The University Of Sheffield.

Downlink Training Sequence Design Based on Achievable Sum Rate Maximisation in FDD Massive MIMO Systems

By

Muntadher Alsabah

A thesis submitted for the degree of

Doctor of Philosophy in Engineering

The University of Sheffield

Faculty of Engineering

Department of Electrical and Electronic Engineering

February 2020

Supervisor

Professor Timothy O'Farrell

Acknowledgements

I would like firstly to thank my first supervisor Professor Timothy O'Farrell for his enormous support, commitment and guidance during my PhD journey, his technical advice and suggestions, subject knowledge and valuable feedback and comments were very beneficial to point me in the correct direction and served to complete the writing and development of this thesis. I also would like to thank my second supervisor Dr Mikko Vehkaperä for his tremendous help, valuable comments and support during my PhD study. His technical suggestions and insightful feedback helped me to develop this thesis. I am indebted to both of my supervisors.

I would like to extend my gratitude to my friends in the communications group that stood with me during my PhD journey. Spacial thanks go to my fellow friends Dr. Stephen Henthorn, Dr. Mohammad Eissa, Dr. Ahmed Al-baidhani and Dr. Baraa Al-Azzawi in the communications group for their support during my PhD journey.

I would also like to thank Mrs Hilary Levesley and Mrs Sarah woods and Mr David Barrott for their enormous help and cooperation during my study.

I would also like to thank my father, mother and sisters for their continued support, help and encouragement throughout my study.

Special thanks go to my wife and children for their patient and support during my PhD journey.

I would also like thank Dr. Azet sadik and Dr. Aws sadik for their support and standing with me during my PhD journey. Special thanks go to my best friend Dr. Hayder Ali Faris. I am very grateful for his help and support.

Finally, I would like to extend my great thanks to Iraqi cultural attache in london and the council of representatives of Iraq.

Abstract

This thesis addresses the key technical challenges related to the design of the downlink (DL) training sequence for the channel state information (CSI) estimation in frequency division duplex (FDD) massive multiple-input multiple-output (massive MIMO) systems with single-stage precoding and limited coherence time. To this end, a computationally feasible solutions for designing the DL training sequences are proposed and novel closed-form solutions for the optimum pilot length that maximises the sum rate with single-stage precoding and limited coherence time are derived. The results in this thesis show that for practical base station (BS) array sizes of $N < 250$ antennas and limited coherence time, the sum rate of an FDD system using DL channel estimation is comparable to the performance of a time division duplex (TDD) system. The results demonstrate that for array sizes of $N > 50$ the diversity of spatial correlations between multiple users achieved more than 40 bits/s/Hz improvement in the sum rate of the regularised zero forcing (RZF) precoder in comparison to uncorrelated channels with identical channel covariance matrices. Finally, the analyses of the complexity results in this thesis show that more than four orders-of-magnitude reduction in the computational complexity is achieved using the superposition design, which signifies the feasibility of this approach for practical implementations compared with state-of-the-art training designs. An asymptotic random matrix theory along with the P-degrees of freedom (P-DoF) channel model are adopted in this thesis to develop an analytical closed-form solution for the sum rate of the beamforming (BF) and RZF precoders, with perfect and imperfect CSI estimation. Excellent agreement between the numerical, analytical and simulated results are obtained, which underpins the contributions of this research. Overall, the proposed approaches open up the possibility for FDD massive MIMO systems operating in a general scenario of single-stage precoding and more realistic channel conditions, particularly channel correlation and limited coherence time.

Contents

List of Figures

List of Tables

List of Abbreviations

4G Fourth generation.

5G Fifth generation.

AoA Angle of arrival.

a.s. Almost surely.

BF Beamforming.

BLAST Bell labs layered space time.

CMI Conditional mutual information.

CSI Channel state information.

CS Compressed sensing.

CSCG Circularly symmetric complex Gaussian.

CGD Conjugate gradient descent.

DFT Discrete Fourier transform.

dB Decibel.

DL Downlink.

DPC Dirty paper coding.

DoF Degrees of freedom.

EVD Eigenvalue decomposition.

EB Exabytes .

FDD Frequency division duplex.

FLOPS Floating point operations per second.

JSDM Joint spatial division and multiplexing.

KKT Karush-Kuhn-Tucker.

LTE long-term evaluation.

MRT maximum ratio transmitter.

MIMO Multiple input multiple output.

Mbit/s/Hz Megabit per second per hertz.

MMSE Minimum mean square error.

mmWave Millimetre wave.

MSE Mean square error.

NLoS Non line of sight.

PCs Personal computers.

PDF Probability density function.

QoS Quality of service.

RMT Random matrix theory.

RF Radio frequency.

RZF Regularised zero forcing.

SNR Signal to noise ratio.

SINR Signal to interference plus noise ratio.

SCMI Sum conditional mutual information.

SMSE Sum mean square error.

SDMA Space division multiple access.

SVD Singular value decomposition.

TDD Time division duplex.

UL Uplink.

ULA Uniform linear array.

UTs User terminals.

UPA Uniform planar array.

List of Notations

$\mathbf{A}$ Matrix.

$\mathbf{a}$ Vector.

$\mathcal{N}$ Normal distribution.

$C\mathcal{N}(\mathbf{0}, \mathbf{R})$ Denotes the circularly symmetric complex Gaussian (CSCG) probability distribution with mean $\mathbf{0}$ and covariance matrix $\mathbf{R}$.

$\mathbb{E}[\cdot]$ Refers to the expectation operator.

$\mathbb{R}$ Set of real numbers.

$\in \{\cdot\}$ A member of set $\{\cdot\}$.

$\mathbb{C}$ Set of complex numbers.

$\mathbb{Z}$ Set of real integers.

$\mathbf{I}_N$ Denotes the $N \times N$ identity matrix.

$\mathrm{tr}(\cdot)$ Trace operator.

$(\cdot)^{\mathrm{T}}$ Transpose operator.

$(\cdot)^{\mathrm{H}}$ Hermitian operator.

$(\cdot)^{*}$ Conjugate operator.

$(\cdot)^{-1}$ Inverse operator.

∞ Infinity.

$\mathrm{diag}(a_1, a_2, \ldots a_N)$ is a diagonal matrix of size N, so that $N \times N$ square matrix with $a_1, a_2, \ldots a_N$ in the main diagonal and zero elements elsewhere.

$\forall$ For all.

$|\cdot|$ Absolute value operator.

$[\mathbf{A}]_{:,j:m}$ Denotes a submatrix containing columns j though m of matrix $\mathbf{A}$.

$[\mathbf{a}]_k$ Denotes the element in the kth row of vector $\mathbf{a}$.

$[\mathbf{A}]_{k,l}$ Denotes the element in the kth row and lth column of matrix $\mathbf{A}$.

$\hat{\mathbf{a}}$ Estimate of a vector $\mathbf{a}$.

$(\mathbf{A} \otimes \mathbf{B})$ Denotes the Kronecker product between matrix $\mathbf{A}$ and $\mathbf{B}$.

$\lceil \cdot \rceil$ Denotes the ceil operation.

$\lfloor \cdot \rfloor$ Denotes the floor operation.

Chapter 1

Introduction

The global mobile data traffic has increased significantly during recent years. Inspection of recent statistics reveal that over the past few years, global mobile data traffic has grown exponentially and it is expected to grow to 77 Exabytes (EB)[1] per month by 2022, i.e. increasing over 7-fold compared with the year 2017 [1]. Fig. 1.1 shows the estimated global mobile data traffic in EB per month [1]. The main drivers of the dramatic growth in the data traffic are personal computers (PCs), laptops, tablets, and smartphones. According to a Cisco report [1], by the year 2022, there will be 8.4 billion mobile-ready devices connected to the Internet. These devices are typically consuming data traffic (mainly video) rather than voice traffic. As a result of such enormous data traffic growth, the current fourth generation (4G) of cellular systems, known as the long-term evolution advanced (LTE-A) networks, will achieve their capacity limits. For example, the theoretical maximum data rate that the current LTE network can achieve with a 2×2 multiple-input multiple-output (MIMO) system is about 150 Mbit/s [3]. Hence, the data rate performance that is currently supplied by 4G networks is unlikely to sustain the enormous ongoing data traffic explosion [4]. This has motivated research into continuing to advance the existing networks, in addition to innovation around new physical-layer techniques for the next generation cellular systems to fulfill the increasing demands for data traffic.

[1] One-exabyte is equivalent to one billion gigabytes.

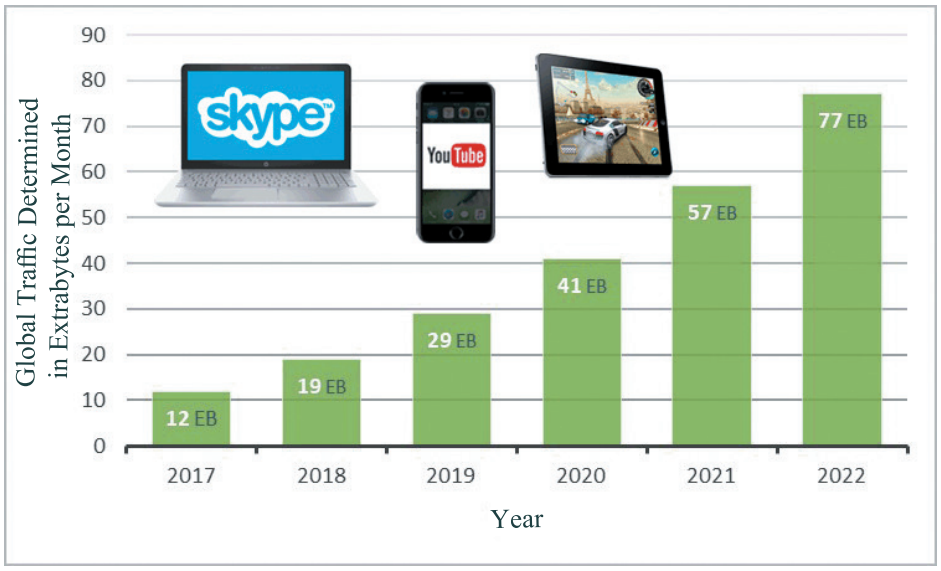


Figure 1.1 Global mobile data traffic demand in EB per month between 2017 and 2022 [1].

1.1 The next generation wireless communication aims and key technology

In order to support massive numbers of connected devices with diverse data rates, latency and energy efficiency requirements, the next generation of mobile communication systems, i.e fifth generation (5G) networks, are being developed [5, 6]. These future mobile networks also aim at: (a) allowing several orders of magnitude enhancement in achievable throughput; (b) achieving a considerable improvement in users' quality of service (QoS); and (c) enabling a significant enhancement in communication reliability [3, 7]. As such, there is an essential requirement to develop a new advanced physical-layer technology to meet the diverse set of requirements of the 5G networks with low-complexity signal processing. Massive MIMO, proposed in [8], is introduced as one of the most promising technologies to achieve this goal. Massive MIMO is a scaled-up version of the conventional space division multiple access (SDMA) system with a large number of base station (BS) antennas giving adaptive beamforming gains and more spatial multiplexing [9, 10]. Specifically, when a large number of antenna elements are deployed by the BS, a significant improvement in the sum spectral efficiency (SE) per-user terminal (UT) and per cell can be achieved. In addition, massive MIMO transmission has several advantages such

as: (a) allowing the use of linear precoding schemes with low complexity signal processing and hardware power consumption; (b) achieving a uniform QoS across the entire cell; (c) providing immunity against fading; and (d) reducing the BS energy consumption [11–13]. These key advantages promote massive MIMO to be the core technology for the future 5G networks [14].

1.2 Aims of the Research and Motivation

1.2.1 Aims

This research aimed to provide an in-depth study of the technical/design challenges with respect to using downlink training sequences for CSI estimation in a frequency division duplex (FDD) massive MIMO system. Therefore, the author of this thesis aimed to address these challenges and explore feasible solutions for the downlink training sequences with a lowest possible level of design complexity. This also included investigating the impact of the proposed training solutions on the achievable sum rate performance of an FDD massive MIMO system with limited coherence time.

1.2.2 Motivation

As explained in Section 1.1, massive MIMO is key for the next generation cellular systems due to its capability to improve the communication reliability and support the enormous growth in data rate demands. The potential of massive MIMO systems has motivated the author of this thesis to investigate its achievable sum rate performance in realistic scenarios. From an information theoretic point of view, the performance of massive MIMO systems depend strongly on the accuracy of the channel state information (CSI) estimation. This thesis focuses on FDD systems, as the CSI can only be obtained by a dedicated training sequence in the downlink. However, the expansion of the number of transmit antenna elements at the base station in massive MIMO makes the downlink training overhead overwhelming in FDD systems, and thus, CSI estimation becomes very challenging with limited channel coherence time. This is because a large percentage of the available coherence time would be dedicated to downlink channel training, and this in turn would directly impact the achievable sum rate. Therefore, feasible

solutions with low computational complexity and minimal overhead for downlink training sequence are crucial in systems with very large non-conventional transmit BS arrays and limited channel coherence time. This has motivated the author of this thesis to address the key technical/design challenges related to the design of the downlink training sequence for CSI estimation in order to enable the FDD mode of operation, which still dominates current cellular networks, to operate over channels with a finite coherence time in massive MIMO systems. Furthermore, providing a closed-form solution for the optimum training sequence length that maximises the achievable sum rate with limited coherence time is essential. To the best of the author's knowledge, no analytical solution for the optimum training sequence length has yet been developed in an FDD based massive MIMO due to the technical challenge of deriving a closed-form. The author of this thesis was motivated to overcome this challenge by developing an analytical closed-form solution, which forms a key contribution of this research. Overall, the proposed design paradigms presented in this thesis have the potential to provide a pragmatic downlink training sequence for CSI estimation in an FDD massive MIMO system, as well as achieve a significant reduction in the computational complexity.

1.3 Thesis Objectives

In order to achieve the research aims, the following set of objectives have been identified:

1. **Design and model an FDD massive MIMO transmission system:** This involved investigating the fundamentals of FDD transmission based on massive MIMO systems and developing a physical layer model using the MATLAB software simulator. To this end, a massive MIMO system model that characterises the achievable sum rate performance based on two linear precoders, the beamforming (BF) precoder and the regularised zero forcing (RZF) precoder is established. To demonstrate the impact of spatial correlation on the downlink CSI estimation of an FDD massive MIMO system, and hence the sum rate, two channel correlation models, namely, the analytical P-degrees of freedom (P-DoF) channel model and the more realistic one ring (OR) channel model are introduced. This research also included investigating the impact of a rich scattering environment modelled

by the uncorrelated Rayleigh fading channels on the achievable sum rate and compared its performance with those of correlated fading channels in an FDD massive MIMO system.

2. **Develop a new gradient-descent iterative algorithm for downlink training sequence optimisation:** This research involved investigating a downlink training sequence optimisation for an accurate CSI estimation in an FDD multiuser massive MIMO system when users exhibit different spatial correlations. The existing gradient-based iterative algorithms that have been proposed in the literature for the training sequence optimisation exhibit slow convergence rates. Therefore, the objective of this research was to develop a new iterative algorithm, which is able to provide a fast convergence speed compared with the existing iterative algorithms. This included analysing the computational complexity of the proposed iterative algorithm and comparing its performance with the state-of-the-art iterative algorithms for training sequence optimisation.

3. **Develop a tractable framework analysis for the optimum downlink training sequence design in FDD massive MIMO systems:** The objective of this research was to establish a tractable framework analysis for the optimum downlink training sequence in the special scenario where users exhibit the same spatial channel correlation. This included providing closed-form solutions for the optimum training structure and the optimum pilot duration with an objective function based on maximising the achievable sum rate in an FDD massive MIMO system. The following steps explain how the investigation was approached:

    - Design an optimum downlink training structure based on sequences constructed from the eigenvectors of the transmit correlation matrix, which are common for all the users. Based on this pilot design, the minimum-mean square-error (MMSE) channel estimation is simplified and a tractable framework analysis is established.
    - Provide an analytical closed-form solution for the optimum pilot sequence length that maximises the achievable sum rate with limited coherence time.
    - Develop a tractable framework analysis based on random matrix theory methods for the achievable sum rate of the BF and RZF precoders with perfect and imperfect CSI estimation in an FDD massive MIMO system with common correlation

between users. This allows analytical solution for the achievable sum rate optimisation problem to be obtained and the numerical evaluation to be readily calculated without resorting to computationally demanding Monte Carlo simulations.

- Analyse the computational complexity of the training sequence design based on a common correlation between users and compare its performance with the state-of-the-art iterative algorithms.

4. **Develop a non-iterative training sequence design for FDD massive MIMO systems in the more general scenario where users exhibit distinct spatial correlations:** The objective of this research was to develop an alternative approach to iterative design for the downlink training sequence, which is feasible with limited coherence time. Building on the analysis in Chapter 3, the performance of the newly developed training design is characterised numerically, analytically and through simulation with the precoding type considered. The following investigations have been performed:

   - Develop a low-complexity non-iterative training sequence design to address the challenge of having different spatial channel correlations for the users in the desired single-stage precoding. This was achieved based on the principle of linear superposition constructed from the user's channel correlation matrices.
   - Characterise the achievable sum rate performance of the proposed superposition training sequence and compare its performance with the state-of-the-art sequence designs based on iterative algorithms.
   - Assess the achievable sum rate performance of the superposition training sequence solution for FDD massive MIMO systems in a realistic channel model and compare its performance with the uplink channel estimation as used in a conventional TDD massive MIMO system.
   - Develop an analytical closed-form solution for the signal-to-interference-plus-noise ratio (SINR), and hence the sum rate, of the BF and RZF precoders with perfect and imperfect CSI estimation in the general scenario where users exhibit distinct spatial channel correlations. This allows the achievable sum rate to be straightforwardly evaluated without executing extensive Monte Carlo simulations.

- Develop an analytical closed-form solution for the optimum pilot sequence length for the general scenario of FDD multiuser massive MIMO systems with different spatial correlations and investigate its suitability in a realistic channel model.
- Analyse the computational complexity of the proposed superposition sequence design and compare its performance against the state-of-the-art iterative sequence designs.

5. **Enhancing the performance of the superposition training sequence design:** This research dealt with a non-uniform power allocation between the data and training phases in order to further enhance the FDD massive MIMO system performance. In addition, the objective of this research was also to introduce a low-complexity weighted superposition training sequence design with the aim to further improve the performance of the superposition design and thus enhancing the achievable sum rate FDD massive MIMO systems.

6. **Explore the linear superposition training design for a uniform planar array:**

   The objective of this research was to investigate if the low-complexity superposition training framework works in another type of array configuration based on the uniform planar array (UPA). This research also included analysing the sum rate of the proposed training scheme and conducted a comparison with the state-of-the-art iterative training designs.

The following section presents a review of the DL training sequence designs relevant to the research in this thesis.

1.4 Review of training sequence designs for massive MIMO

This section first presents a review of training sequences in the conventional point-to-point MIMO systems. Then, the state-of-the-art training sequences that have been proposed for addressing the challenge of DL channel estimation in the FDD massive MIMO systems are reviewed.

From the early generation of wireless communications systems, the training sequence design and channel estimation attracted attention due to their essential role in precoding and decoding designs, interference suppression, resource allocation, rate adaptation, and improving the communication link reliability [15]. Several studies have investigated the design of training sequences in conventional MIMO systems with an objective function based on minimising the mean squared error (MSE) criterion of the channel estimates. In particular, the works in [16–24] have found that a significant improvement in the CSI estimation performance can be achieved by utilising Bayesian estimation [25], i.e. adopting the minimum mean square error (MMSE) filters, which make use of channel and noise statistics. The research in [26] investigated the effect of DL training sequence and channel estimation on the capacity of the conventional point-to-point MIMO systems in uncorrelated Rayleigh fading channels. A conclusion arising from [26] is that; to achieve a decrease in the channel estimation error, the DL training sequence length should be scaled linearly with the number of BS antennas. Application of this design principle to the DL training sequence and channel estimation of an FDD based massive MIMO with limited coherence time, as discussed in [12, 27, 28], is indeed unfeasible for large N, thereby rendering DL training unsuitable for FDD operation in an uncorrelated channel. The argument of the present thesis is that following this design paradigm leads to unnecessarily pessimistic predictions on the performance of the FDD massive MIMO systems with DL training.

Early research on massive MIMO systems focused on the TDD operation, discussed in subsection 2.3.1, where the required CSI can be obtained by sending a superposition of orthogonal sequences over a length of T_{p} symbols in the UL direction during each coherence interval [29–31]. The authors in [32, 33] found that the number of UL training symbols is proportional to the number of users K and independent of the number of BS antennas N that can be made as large as required. However, despite the promising results of TDD operation with massive MIMO systems [34], with partial channel estimation the transceiver hardware impairments and calibration error in the UL/DL RF chains of the TDD based systems are generally limiting factors [35–38]. Besides, the signal-to-noise-ratio (SNR) in TDD systems is typically lower than in an FDD based system because the power amplifier is enabled only part of the time [39]. In addition, over 85% of the current commercial long term evolution (LTE) wireless mobile networks operate in FDD mode [40]. Therefore, opening up the possibility for realising FDD operation

in massive MIMO systems with limited coherence time is of interest.

To address the challenge of FDD massive MIMO systems, several studies have investigated the design of training sequences using different channel models and design criteria. The salient research studies of training sequence designs in the FDD massive MIMO systems can be divided into three categories; spatial and temporal correlations based, compressive sensing based, and spatial correlation based. Research directions that incorporate correlations both in the time and spatial domains, which are referred to as a spatio-temporal channel correlations, have been considered in [41–43]. These works explored the joint use of the spatial and temporal channel correlations. Kalman-filter is used to track the channel variation and to improve the DL channel estimation performance. Specifically, the BS transmits the DL training signal to the UT, and the UT estimates the current CSI using an MMSE channel estimate. Then, the UT transmits the best training signal that achieves a minimum MSE performance of the channel estimate back to the BS. The BS would then identify the best channel state for the current time instance and use the temporal channel correlation based on a Gauss-Markov process with Kalman-filtering to predict the CSI of the next time instance through the use of a Greedy sequential iterative algorithm. In these investigations, the sequential iterative training sequences and the Kalman-filtering are beneficial to overcome the overhead of DL training and minimise the MSE of the channel estimate in an FDD based massive MIMO system. The training sequences are designed based on the criterion of minimising the MSE of the channel estimate and maximising the received SNR for a massive MIMO system in a scenario where all users exhibit a common spatial correlation. However, minimising the MSE of the channel estimate maximises the received SNR term while affecting the pre-log fraction term $(1 - T_{\mathrm{p}}/T_{\mathrm{c}})$. Therefore, to maximise the sum rate, both terms should be taken into account.

Besides the research studies that exploited the spatio-temporal correlations, another line of research studies have focused on the design of training sequences by utilising compressed sensing (CS) based techniques [44–47]. The sparsity structure on the virtual angular domain is exploited to help constrain the training sequence length required for the DL channel estimation of an FDD based system. In these investigations, the DL channel estimation problem has been reformulated as a CS problem to reduce the amount of overhead required for channel training. To this end, CS based algorithms, which need a prior knowledge of the channel sparsity, are

used for the DL channel estimation. Channel sparsity can be realised when the ULA at the BS has a large number of antenna elements for a limited number of scatterers, where the narrow angular spread of the outgoing rays at the BS results in high correlations between distinct paths that link the BS and mobile user [48, 49]. While these investigations overcome, at least in part, the infeasibility of DL training in FDD massive MIMO systems, the algorithms used for channel estimation are computationally complex [50]. In addition, due to the unknown sparsity nature of the channels, the CS based approaches cannot be applied in practice to predict the optimum training sequence length that maximises the DL achievable sum rate of an FDD massive MIMO system.

Another development for the DL training sequence and channel estimation in an FDD massive MIMO systems, which goes beyond the temporal correlations and the CS based approaches, is joint spatial division and multiplexing (JSDM) [51, 52]. In JSDM, the mobile users are spatially partitioned into distinct groups and a two-stage precoding process is performed. The role of JSDM is to exploit correlations in the spatial domain, where the mobile users within each group exhibit the same spatial correlation. Specifically, in the first-stage precoding (pre-beamforming), the precoder is designed based on the group of users that have similar channel covariance matrix, where the corresponding eigenvectors are assumed to be the same for all the users within the same group. In the second-stage precoding, the BS would then need to use a classical multiuser precoding based on instantaneous CSI to control the residual interference between the users in different groups. The research in [51, 52] uses the sum of the eigenvectors of each group to perform the first of two stages of precoding, essentially forming a beam for each group. This summation can be considered a linear superposition. The authors found that, when using a linear superposition pilot framework with a users grouping-based approach, the training sequence length in the DL can be scaled linearly with the number of user groups, which can be less than N, resulting in feasible pilot overhead requirement for FDD operation.

Though a two-stage precoding-based approach helps to constrain the DL training sequence length for CSI estimation using the principle of linear superposition, opportunistic scheduling and clustering algorithms of the user groups, and of the users inside each group, are essential for the JSDM approach to work. Otherwise, in realistic propagation conditions with heterogeneous user channels, the system performance could be significantly degraded due to the residual

interference between the users in the same group and the users in different groups, considering a finite number of mobile users to be selected at the scheduling-stage. In addition, the research in [51, 52] does not address the single-stage precoding scenario with K independent user covariance matrices, and thus, cannot predict the optimum pilot sequence length that maximises the achievable sum rate with limited coherence time in this preferred case.

To address the challenge of having different spatial channel correlations for the users in the desired single-stage precoding, several studies have investigated the design of training sequences by considering different optimisation algorithms, see e.g., [53–56]. For example, the work in [55] considers spatially correlated multiuser massive MIMO systems and the DL training sequence is investigated via the use of an iterative algorithm with the criterion of minimising the MSE of channel estimates. An information theoretic DL training sequence design and channel estimation for an FDD multiuser massive MIMO system is considered in [54]. There the DL training sequence is designed based on maximising the sum of conditional mutual information (SCMI) between the mobile users channel vectors and their corresponding channel estimates. The DL training sequence is optimised using an iterative gradient-based algorithm on the complex Grassmannian manifold[2] in the presence of spatially correlated fading and multiple Gaussian mixture distributions. In [53], a DL training sequence of an FDD multiuser massive MIMO system is optimised by utilising a Karush-Kuhn-Tucker (KKT) iterative gradient-based algorithm. To this end, in the presence of spatially correlated fading with white noise and a smooth Gaussian variable, the DL training sequence is found iteratively as a solution to a sum SCMI maximisation problem. Similar to the system model assumption in [53], the authors in [56] investigated the design of DL training sequence of an FDD multiuser massive MIMO system by considering a sum mean square error (SMSE) minimisation problem. Using a SMSE optimisation criteria, the training sequence is obtained through the use of an iterative gradient descent algorithm on the complex Grassmannian manifold. Note that the design principles in [53–56] are more general compared with [51, 52] since they do not enforce a two-stage precoding structure for the DL channel estimation in FDD massive MIMO systems.

The shared conclusion arising from [53–56] is that when the mobile users admit different

[2]In mathematics, the complex Grassmannian manifold of a space is defined as a set of reduced dimension subspaces from the whole space, which can be used to solve different optimization problems [57].

spatial channel correlations, iterative algorithms are required to optimise the training sequences since they span distinct channel covariance matrices. However, despite the attractiveness of iterative solutions when it comes to optimising the DL training sequences in FDD multiuser massive MIMO systems, these solutions still overly complex, making practical implementation problematic. This is because the algorithms require many iterations to reach convergence especially for $N, K \gg 1$. Each iteration includes the inversion and multiplication of large matrices, which requires a lot of computing resources. As such, the convergence time of these algorithms are high, making such algorithms unsuitable for practical systems. Further, in the case of distinct spatial correlations, finding the optimum training sequence length is very challenging due to the technical challenge of deriving closed-form solutions based on such techniques. Nonetheless, such iterative optimisation approaches provide useful indication of the best sum rates, which more practical solutions, as developed in this thesis, can aim for.

1.5 Thesis Contributions and Originality

This thesis addresses the challenges of downlink channel estimation using low computational complexity training sequences in an FDD massive MIMO communication system. This also includes providing important insights and recommendations into the practical perspective of an FDD massive MIMO system. As an outcome of this research, the following original contributions are carried out in this thesis:

1. The proposal of new computationally efficient conjugate gradient-descent iterative algorithm based on the Riemannian manifold in order to optimise the downlink training sequence in an FDD multiuser massive MIMO system when users exhibit different spatial correlations. The analyses of the results show that the proposed iterative algorithm achieves a robust sum rate performance owing to a faster convergence rate.

2. The development of the simplified framework analysis for the optimum training sequence structure and the optimum training duration based on a scenario where all users exhibit the same spatial correlation at the BS. This development is essential to provide the full solution for the downlink training sequence design in FDD massive MIMO system.

3. The proposal of a non-iterative training sequence design for the more general scenario where the users exhibit different spatial correlations, which is feasible with limited coherence time. This is achieved based on the principle of linear superposition of sequences constructed from the eigenvectors of K independent users' channel correlation matrices. This study shows that the proposed training sequence design achieves almost the same rate performances as state-of-the-art training designs, while reducing the computational complexity significantly. Furthermore, the superposition sequence design is evaluated analytically, numerically and through simulation, which demonstrates excellent agreement throughout (see Chapter 5).

4. The straightforward (explicit) mathematical analyses of the achievable sum rate for the BF and RZF precoders in FDD massive MIMO systems with perfect and imperfect CSI estimation. These analyses are developed based on the random matrix theory methods and the P-DoF channel model. The rigorous mathematical analyses tightly agree with the simulated results. These analyses are used to avoid executing extensive Monte Carlo simulations, which allows a significant reductions in the computational complexity when designing a training sequence.

5. The new closed-form analytical solution for the optimum downlink training sequence length that maximises the achievable sum rate of the BF precoder in an FDD massive MIMO system with distinct spatial correlations and limited coherence time. The analyses of the results show that the optimum pilot length that is analytically optimised for the BF precoder can also be used to predict the sum rate performance of the RZF precoder. Furthermore, the numerical results illustrate that this pilot framework design is also sufficient to predict the sum rate performance of the BF and RZF precoders in the more realistic channel models. This finding results in large reductions in the computational complexity.

6. The computational complexity analyses of various downlink training sequence solutions. These analyses are compared against each other in this thesis in order to show the best deployment options that are able to deliver a feasible sum rate performance at a lower complexity.

7. The investigation of the non-uniform power allocation between the training and data

phases and evaluating its impact on the achievable sum rate performance in an FDD massive MIMO system.

8. The proposal of a new weighted superposition training sequence design to improve the performance of superposition pilot scheme which will result in further sum rate enhancement in an FDD massive MIMO system with limited coherence time, especially with relatively weak correlated channels.

9. The design and evaluation of a low computational complexity superposition training design for the uniform planar array.

The aforementioned contributions allow feasible solutions for downlink training sequence design and CSI estimation in an FDD massive MIMO system to be achieved with a significant computational complexity reduction. In particular, the proposed training solutions open up the possibility for FDD massive MIMO systems operating in a general scenario of single-stage precoding and distinct spatial correlations with limited coherence time.

1.5.1 List of publications

[1] M.Alsabah, M. Vehkapera, and T. O'Farrell, "Non-iterative Downlink Training Sequence Design Based on Sum Rate Maximization in FDD Massive MIMO Systems," IEEE Access, vol. 8, pp. 108 731–108 747, May 2020.

1.6 Thesis outline

This thesis is structured into seven chapters, a key summary of each chapter development is provided below:

- Chapter 1 presents a brief introduction about next generation cellular systems and explains briefly why new physical layer techniques are required. The aims for this thesis and the research motivation are then provided. This chapter also summarises the planned objectives of the research. The areas of novelty and originality of research contributions

are also detailed. Chapter 1 reviewed different approaches for downlink training sequence design and explaining briefly the key properties of each approach. Chapter 1 ends by reviewing the state-of-art research undertaken to address the challenge of FDD operation in massive MIMO systems, which sets the space for the contribution of this thesis. Overall, Chapter 1 demonstrates the potential for a low computational complexity training approach and emphasises the necessity of the research carried out in this thesis. Finally, some useful definitions are outlined at the end of this chapter.

- In Chapter 2, an overview of massive MIMO communications system is provided, which includes a general introduction to the research topic. Chapter 2 introduces the necessary background materials that are relevant to the thesis.

- Chapter 3 first introduces the downlink massive MIMO system model based on linear precoding, which will be then used in this thesis. In addition, Chapter 3 proposes a conjugate gradient descent iterative algorithm to optimise the downlink training sequence in an FDD massive MIMO system when the users exhibit distinct spatial channel correlations while improving the convergence speed compared with the state-of-the-art iterative algorithms. The computational complexity analysis of the proposed iterative algorithm is also characterised in Chapter 3 and compared with the state-of-the-art iterative algorithms. Furthermore, the achievable sum rate performance of the proposed iterative approach for downlink sequence optimisation is analysed and compared with the rates achieved by the state-of-the-art iterative algorithms.

- Chapter 4 focuses on establishing a tractable framework analysis for the optimum training sequence in an FDD massive MIMO system based on a spacial scenario where users exhibit common spatial correlations. To this end, an optimum pilot sequence design is developed and closed-form solution for the optimum training sequence that maximises the achievable sum rate with limited finite channel coherence time is provided. The computational complexity of the training sequence design is also quantified and compared with the state-of-the-art iterative approaches.

- Building on the framework analysis and theoretical development provided in Chapter 4, Chapter 5 investigates a non-iterative sequence design for CSI estimation in the more gen-

eral scenario where the users have different spatial correlations. To this end, the principle of linear superposition of sequences constructed from the eigenvectors of distinct users' channel correlation matrices is explored. This allows a feasible solution for the problem of downlink CSI estimation in an FDD massive MIMO system to be achieved with a significant reduction in the computational complexity. Based on the superposition training design, an analytical closed-form solution for the optimum training sequence length is provided. Furthermore, the computational complexity of the proposed superposition sequence design is analysed and compared with the state-of-the-art training sequence designs.

- In Chapter 6, a comprehensive research investigation is conducted in order to find new possible low-complexity approaches to enhance the downlink CSI estimation in an FDD massive MIMO system. In particular, the research in Chapter 6 includes investigating the achievable sum rate optimisation using a non-uniform power allocation between the data and pilot signals. Furthermore, based on the investigation in Chapter 5, a new weighted superposition training sequence design is proposed in Chapter 6 aiming to further enhance the superposition approach (i.e. unweighted) developed in Chapter 5, and thus, improve the achievable sum rate performance. Chapter 6 ends by an investigation of whether the low-complexity linear superposition pilot design is also applicable to the uniform planar array.

- Finally, the research outcomes and concluding remarks are summarised in Chapter 7. In addition, useful recommendations for the future work directions are also outlined at the end of Chapter 7.

  Fig. 1.2 demonstrates a summary of main contributions in each technical chapter.

1.7 Useful definitions

CSI estimation: In wireless communications systems, CSI estimation is the instantaneous knowledge that the BS has of channel properties of a communication link. In par-

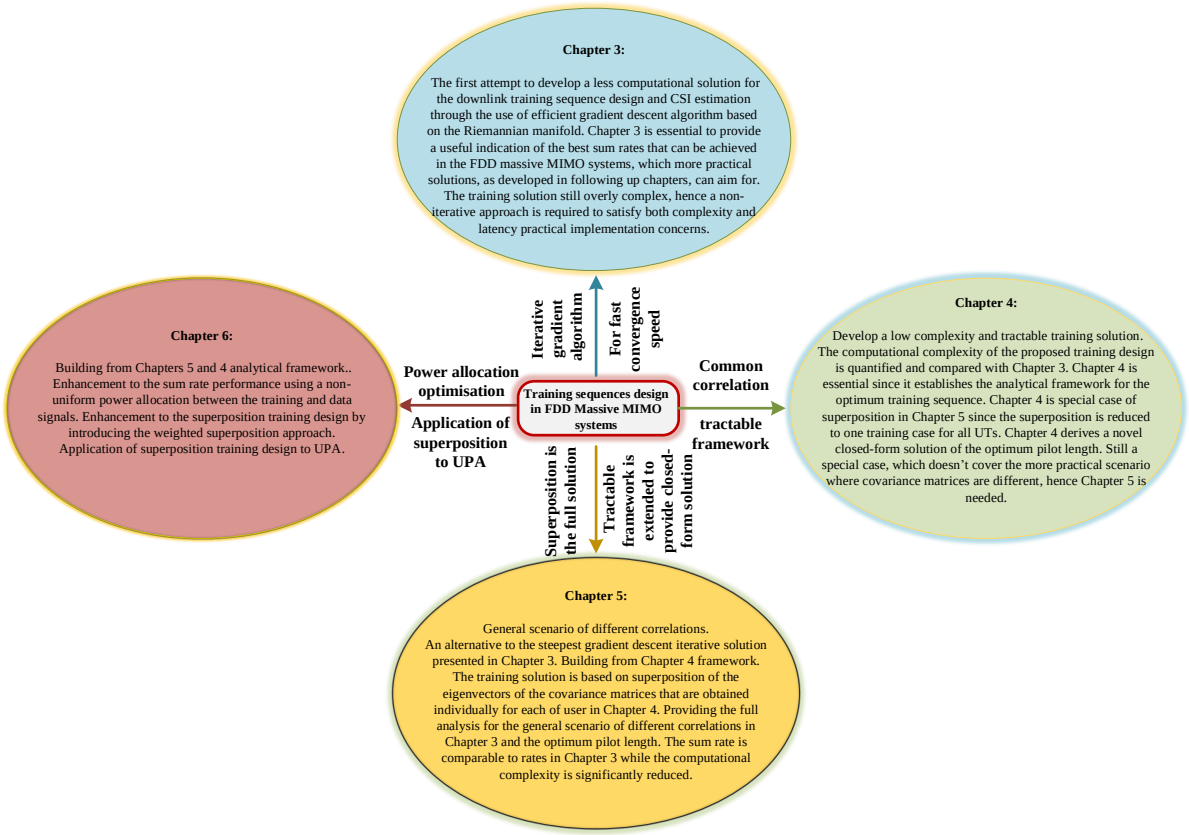


Figure 1.2 Summary of the key contributions per each technical chapter.

ticular, CSI estimation is an essential process that is required to characterise the channel and provide an information of how a signal propagates from the BS to the UTs.

Favourable propagation: means that the channel matrix between the BS and the UTs is well-condition, which is held in the massive MIMO systems due to the low of large number. As such, the channel vector of different UTs become orthogonal as N increases, and thus resulting in mitigating the inter-user interference.

Channel hardening: means that the effective channel after precoding/beamforming is almost constant, i.e., the communication link behaves as if there is no small-scale fading.

Spatial correlation: When the BS antenna elements are closely spaced and the mobile users experience limited scattering geometries, the channels become strongly spatialy correlated. Specifically, the spatial correlation depends mainly on the antenna array geometry (aperture, antenna spacing) and the angular spread, and must be accounted for by any realistic channel model.

Common spatial correlation: means that all the users experience the same spatial correlation at the BS.

Distinct spatial correlations: means that each user experiences different spatial correlations at the BS.

The coherence time: is the time interval over which the channel responses, i.e. the phase and amplitude variations due to UT mobility are approximately constant. The coherence time is inversely proportional to the carrier frequency.

The coherence bandwidth: is the frequency interval over which the channel responses are approximately constant. The coherence bandwidth is inversely proportional to the delay spread.

Delay spread: In wireless communications systems, the delay spread represents the time difference between the earliest multipath component and last multipath components.

SINR: The SINR is defined as the power of a certain signal of interest divided by the sum of the interference power, which introduces from all the other interfering signals, and the noise power.

Precoding: is a digital signal processing technique, which is used by the BS to transmit one or multiple spatially directive signals to spatially multiplex many users. The precoding technique is able to direct the signal into a specific user direction where the phases and powers can be jointly designed to achieve high sum rate.

BF precoding: is a linear signal processing technique, which can be obtained through the conjugate transpose of the DL channel matrix, i.e. $\mathbf{H}^{\mathrm{H}}$. The BF precoder is also known as maximum ratio transmission (MRT), which maximises the signal power that is intended to the UTs. The BF precoder is a low complexity processing technique since it does not require a matrix inversion.

RZF precoding: is a linear signal processing technique that transmits the signal toward the desired UT and eliminates the interference caused by other UTs. RZF is considered as the state-of-the-art precoder scheme for the MIMO wireless communications systems due to its capability of trading off the advantages of the BF precoder and the well known Zero-forcing (ZF) precoding. As such, the achievable sum rate would be maximised using the

RZF precoder. At high SNR values, the interference is increased, so the RZF precoding has the ability to suppress the interference while the BF precoder does not. However, unlike the BF precoder, RZF precoding technique involves a matrix inversion.

Monte Carlo simulations: are defined as a general class of computational algorithms, which depends on repeated random sampling in order to obtain numerical results.

Closed-form solutions: mean simplified analytical results, which are obtained without the use of Monte Carlo simulations.

Manifold: In mathematics, the manifold of a parameter space is defined as a set of reduced dimension subspaces from the whole space, which can be used to solve different optimisation problems.

Gradient descent: is defined as a first-order iterative optimisation algorithm, which is useful to find a local minimum of any differentiable function. In particular, steps proportional to the negative of the gradient should be carried out to find a local minimum of a function using gradient descent.

Hermitian matrix: In mathematics, the Hermitian matrix is a position definite matrix where all its eigenvalues are positive and its i-th row and j-th column is equal to the complex conjugate of the element in the j-th row and i-th column.

SDMA is a transmission technique to spatially multiplexes several multiple users streams simultaneously over the same time and frequency.

Matrix exponential: In mathematics, the matrix exponential is used to solve systems of linear differential equations very fast and efficiently.

1.8 Summary of the chapter

Current wireless networks are unlikely to sustain the enormous ongoing data traffic explosion. To satisfy the exponential growth in data traffic, advanced physical layer techniques are crucial. Massive MIMO systems are a key physical layer technology for the next generation wireless communications required to support the huge increasing demand in the

data traffic and network capacity. However, to achieve the full gain of massive MIMO systems, sufficiently accurate and timely estimates of the CSI are required.

This chapter has presented the aims for this thesis and the research motivation. This chapter has also described the planned objectives of this thesis. The areas of novelty and originality of research contributions have been summarised. In addition, this chapter has reviewed several different approaches for the DL training sequence designs in massive MIMO systems. Several studies that consider the challenge of channel estimation in FDD massive MIMO systems are critiqued. Some approaches have used the principle of linear superposition to obtain a training sequence solution, but these require two-stage precoding and sophisticated algorithms for user groups, constraining the approach. In other approaches with single-stage precoding, the training sequences have been designed iteratively by using gradient-based iterative algorithms. While advanced gradient-based iterative algorithms have been developed for this purpose, they provide no closed-form solutions on either the optimum structure of a DL training sequence, nor on the minimum pilot length that maximises the achievable sum rate with limited coherence time. Furthermore, the limited coherence time interval implies that the CSI estimation should be determined more frequently, hence a tractable training algorithm that exhibits fast convergence speed or without resorting to computationally demanding iterative solutions is preferred. For this reason, there is an essential requirement to develop low-complexity training solutions that are feasible for practical systems. In addition, providing a closed-form solution for the optimum training sequence length that maximises the achievable sum rate when a large number of BS antenna elements and a limited coherence time is crucial. This thesis adopts the principle of linear superposition of K independent channel covariance matrices in the single stage-precoding to find a closed-form solution for the optimum training length when users exhibit distinct spatial correlations.

Notation: In this thesis, an upper boldface symbol stands for a matrix whereas a lower boldface symbol stands for a vector. $\mathcal{CN}(\mathbf{0}, \mathbf{R})$ denotes the circularly symmetric complex Gaussian (CSCG) probability distribution with mean $\mathbf{0}$ and covariance matrix $\mathbf{R}$. The term $\mathbb{E}[\cdot]$ refers to the expectation operator. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. The

operators trace, transpose, Hermitian transpose, inverse and absolute value are denoted by $\mathrm{tr}(\cdot)$, $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, $(\cdot)^{-1}$, and $|\cdot|$, respectively. $[\mathbf{A}]_{:,j:m}$ denotes a submatrix containing columns j though m of matrix $\mathbf{A}$. The notations $[\mathbf{a}]_k$ and $[\mathbf{A}]_{k,l}$ are used to denote the element in the kth row of vector $\mathbf{a}$ and the element in the kth row and lth column of matrix $\mathbf{A}$, respectively.

Chapter 2

Background

2.1 Introduction

Next generation cellular systems aim to maximise spectral efficiency in order to satisfy the rapidly increasing demand for wireless data services [3, 58], whilst reducing both the cost and energy consumed [59, 60]. Massive multiple-input multiple-output (massive MIMO) is proposed to achieve this goal [8].

In this chapter, a brief overview of massive MIMO communication systems is provided, including their potential and the challenges to their use. The channel state information (CSI) is defined and the commonly used duplex operation modes namely; the time division duplex (TDD) and frequency division duplex (FDD) modes; are explained. An introduction to the achievable sum rate metric, which will be considered in later chapters for the system performance evaluation, is provided. The essential mathematical tools that are used in this thesis are briefly introduced. Then, spatial channel correlation, which provides a realistic performance assessment of the massive MIMO systems, is reviewed.

2.2 Massive MIMO systems

Massive MIMO is a large-scale multiuser MIMO technique where the base station (BS) is equipped with a relatively large number of transmit antenna elements N, which serve a much

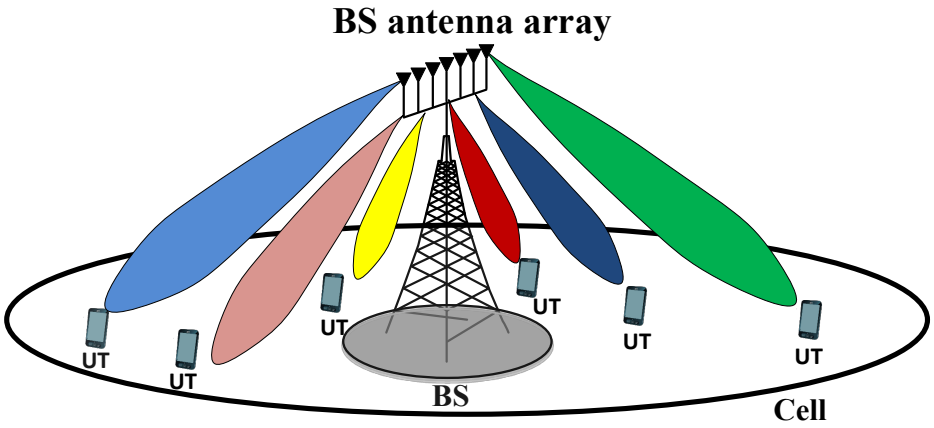


Figure 2.1 BS with massive number of antenna elements that serves simultaneously a K single-antenna UTs.

smaller number of K single-antenna noncooperative user terminals (UTs) [28, 61], as shown in Fig. 2.1. In particular, the more antenna elements that a BS is equipped with, the more degrees of freedom the propagation channel can offer, which translates to an enhanced system performance in terms of data rate and link reliability [13, 62].

2.2.1 Massive MIMO advantages

The deployment of massive MIMO in next generation cellular communications systems is expected to bring a significant enhancement in both spectral and energy efficiencies [11, 31]. Using a large-scale antenna array at the BS provides more distinct paths over which the transmitted signals can propagate, and hence result in increasing the multiplexing gain [61, 63].

In massive MIMO systems, the favourable propagation condition may be achieved in the rich scattering environment, wherein the channel vectors for UTs become asymptotically orthogonal as the number of BS antennas N grows without bound [27]. As such, the effect of uncorrelated noise and fast fading are asymptotically mitigated [8]. Massive MIMO systems also offer a capability to suppress interference, which can be achieved by the spatial resolution of a large-scale antenna array at the BS. In addition, massive MIMO allows a low-complexity digital basedband signal processing techniques to be used in both uplink (UL) and downlink (DL) transmissions [28]. For example, in the DL, linear precoding techniques can be utilised by the BS to focus the transmitted signal to the desired UT, whereas in the UL, linear receive combining schemes can be used at the BS to detect signals transmitted from different UTs. Table 2.1

summarises the key differences between the conventional MIMO and Massive MIMO systems.

In massive MIMO systems, the energy transmitted from the BS can be focused into spatial directions where the UTs are located. When a large number of antenna elements are deployed by the transmit BS, the array gain can be significantly increased. Hence, the radiated power from each of the element can be reduced [11]. Further, since massive MIMO systems enable low-complexity signal processing techniques to be used for precoding and combining, a significant energy saving from the power consumption on signal processing and electrical circuit components can also be achieved [64].

Fig. 2.2 demonstrates a real-time massive MIMO testbed developed by a research group from the University of Bristol and Lund University [2, 65]. The experiment showed that about 145.6 bits/s/Hz sum rate can be achieved using a massive MIMO system with a shared of 20 MHz radio channel at a carrier frequency of 3.51GHz. The results demonstrated that up to 12 UTs can be spatially separated and served over the same time and frequency using a massive MIMO system.

2.2.2 Channel state information

In wireless communication systems, the channel responses that contain the information of the radio wave propagation are typically used for the UL and DL signal processing. In practice, the channel responses need to be estimated regularly for achieving reliable communication systems and efficient resource allocation [66]. The instantaneous knowledge that the BS has of channel response realisations is known as the channel state information (CSI). The statistical CSI information about the distributions of channel responses is assumed to be available anywhere in the network, while instantaneous CSI regarding the current channel realisations require to be estimated at the same pace as the channels change [39]. The availability of accurate CSI allows the BS to adapt transmissions to current channel conditions, and thus, the received signal-to-interference-plus-noise (SINR) can be optimised. Also, the performance of transmit precoding is entirely determined by the CSI accuracy at the BS [67]. Following a stationary random process, the realisations of the channel responses may be changed in every coherence time interval, which are enumerated in symbols per transmission block [39]. In practice, obtaining accurate CSI is typically restricted by how fast the channel realisations are changing with

Table 2.1 Comparison between the conventional MIMO and massive MIMO systems.

The key parameters	**Conventional MIMO systems**	**Massive MIMO systems**
Communication type	Point to point	Point to multiple points
Number of Antenna	$N \leq 8$	$N \geq 16$
The relation between N and K	$N = K$	$N \gg K$
Throughput	Lower throughput	More data rate throughput
Bandwidth	Used more bandwidth	An enhanced usage of the available bandwidth
Degrees of freedom	Not scalable so that provide less degrees of freedom	Provide more degrees of freedom
Degrees of correlation	Typically uncorrelated channel	Highly correlated channel
Coverage area and cell-edge performance	Provide less coverage	Provide more coverage where the cell-edge SNR increases proportionally to N
Fading effect	More sensitive to the fading	Provide immunity against fading
CSI estimation requirement	Less sensitive to the CSI	Sufficiently accurate CSI estimation are required
Heavily loaded cell	Not preferable	Preferable with spatial multiplexing of many users
Reliability	Less reliable	Improve the link reliability with more diversity gain
Energy Efficiency	Less energy efficient	More energy efficient
Antenna coupling effect	Low coupling effect	High coupling effect
Array gain	Less array gain	Provide high array gain
Users service	Typically single user	Multiple users can be served simultaneously over the same time and frequency
Link quality after precoding/combining	Varies over time and frequency, due to frequency-selective and small-scale fading	Almost no variations over time and frequency due to the advantages of channel hardening
Resource allocation	The resource allocation must change rapidly to account for channel variations	The allocation can be planned in advance since the channel quality varies slowly

Figure 2.2 An assembled BS antenna array with massive MIMO testbed [2].

respect to the coherence time interval.

2.2.3 Channel state information estimation

In the typical wireless communications systems, the signal propagates through a medium, which is known as channel. When the signal propagates though the channel, it gets distorted and various noise is added. However, such distortion and noise are required to be removed in order to properly precoder the information data and decode the received signal without significant errors and interference. Specifically, the characteristics of the channel need to be known at the BS. The process to characterise the channel is called CSI estimation. To obtain the CSI at the BS, known training sequences should be transmitted from the BS to the mobile users or vice

versa, depending on the considered duplex operation modes (i.e., either time division duplex (TDD) or frequency division duplex (FDD) modes). The duplex operation modes are further discussed in Section 2.3. The training sequences are predefined signal, which are typically known as "reference signal" or "pilot signal". During the channel estimation process, the CSI is obtained by comparing the transmitted signal and the received signal. Every training signal that is transmitted could have been a signal that carried payload data information, thus the training overhead caused by pilot signaling needs to be minimised. To this end, an optimised channel estimation performance in the DL is achieved by utilising Bayesian estimation, i.e. employing a minimum-mean square-error (MMSE) filter, which makes use of channel and noise statistics.

The training sequences are scaled by the available power and transmitted as signal over a length of T_{p} symbols per each transmission block. In each transmission block, the available energy for transmission can be freely distributed between the training-phase and the data payload.

An example of generating training sequences using a discrete Fourier transform (DFT) matrix is given in (2.1) [68]

$$\mathbf{S}_{\mathrm{p}} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_{T_{\mathrm{p}}} & \omega_{T_{\mathrm{p}}}^{2} & \cdots & \omega_{T_{\mathrm{p}}}^{(T_{\mathrm{p}}-1)} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & \omega_{T_{\mathrm{p}}}^{(T_{\mathrm{p}}-1)} & \omega_{T_{\mathrm{p}}}^{2(T_{\mathrm{p}}-1)} & \ddots & \omega_{T_{\mathrm{p}}}^{(T_{\mathrm{p}}-1)(T_{\mathrm{p}}-1)} \end{bmatrix}, \tag{2.1}$$

where $\omega_{T_{\mathrm{p}}} = e^{-j2\pi/T_{\mathrm{p}}}$ is a T_{p}-th primitive root of 1. The elements of the DFT pilot matrix are located at T_{p} distinct equally spaced points on the unit circle, which corresponds to a T_{p}-ary phase-shift keying (PSK) constellation [39].

2.2.4 Massive MIMO system challenges

The advantages of massive MIMO systems discussed in subsection 2.2.1 are based on the optimistic assumption of an unlimited number of BS antennas. In practice, the propagation environment and the size of the array are limited by a finite physical volume [61]. In addition, limited coherence time is still a major restriction in massive MIMO systems, affecting the length of training sequences required for obtaining accurate CSI [69]. In particular, in contrast to the conventional point-to-point MIMO systems where a small number of antennas N is used at the

BS, in the massive MIMO systems with a large number of antenna elements, the training overhead required for obtaining accurate CSI is potentially overwhelming [27]. This may reduce the spectral efficiency significantly, and thus limit the advantages of massive MIMO systems. Typically, the overall coherence time relies on several factors such as: (a) the delay spread of the propagation environment; (b) the carrier frequency of the communication system; (c) and the mobility of the UTs [70]. In high mobility environments and where high carrier frequencies are used, the channel exhibits a shorter coherence time, so CSI estimation must take place more often. Consequently, in order to carry out CSI estimation within a limited coherence time, the number of training symbols used needs to be reduced. However, using fewer training sequences might degrade the CSI estimation accuracy significantly.

Typically, the channel statistics remain constant over the coherence block length. However, the channels fade independently from symbol block to block. Each coherence transmission block consists of a number of time and frequency resources. In particular, the coherence block consists of a number of subcarriers in frequency (i.e., coherence bandwidth β_c, which represents the frequency interval over which the channel responses are approximately constant) and time samples τ_c, which represents the time interval over which the channel responses can be approximately constant [39]. In general, the channel coherence block length in symbols ranges from hundreds (e.g., in high mobility environments and with high carrier frequencies) to thousands of samples (e.g., in low- mobility environments and with relatively lower carrier frequencies). The number of samples in a coherence block of propagation channels, which corresponds to resource transmission in the time and frequency plane, is obtained based on the Nyquist- Shannon sampling theorem as $T_c = \beta_c \times \tau_c$. In what follow, an example is provided, which demonstrates how to determine the coherence block length in symbols per each transmission block.

In an outdoor environments with τ_c = 1ms and β_c = 200kHz, a UT mobility of v = 37.5m/s = 135km/h can be supported. As such, the coherence block length in symbols in such high mobility environments would be T_c = 200 [39]. In an indoor environments with τ_c = 50ms and β_c = 1MHz, a UT mobility of v = 0.75m/s = 2.7km/h can be supported. The coherence block in symbols in such low mobility environments would be T_c = 50 000 [39].

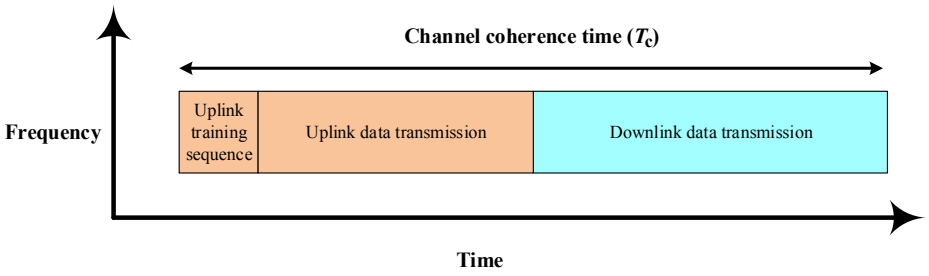


Figure 2.3 Slot structure of the block fading model in TDD operation mode.

2.3 Duplex operation modes

The TDD and FDD are commonly used duplex operation modes in wireless communications systems. The following subsections explain how each method obtains the CSI at the BS in massive MIMO systems.

2.3.1 TDD mode in massive MIMO systems

In TDD based systems, both the UL and DL channels occupy the same frequency band and operate in turns. The underlying assumption in TDD systems is that the overall equivalent complex baseband channel, also accounting for the radio frequency (RF) chains at both ends of the communication link, is reciprocal. In TDD based systems, to obtain the CSI at the BS, the mobile users send mutually orthogonal training sequences, of length K symbols per transmission block, simultaneously to the BS on the UL transmission slots. The BS estimates the UL channel based on the received training signals. Under UL and DL channel reciprocity that holds in TDD systems, the UL channel estimates are used for designing the DL precoder at the BS without the requirement for DL channel estimation [71, 72]. Fig. 2.3 demonstrates the basic block length with a slot structure for TDD operation.

2.3.2 FDD mode with massive MIMO systems

In non-reciprocal channels, such as those encountered in frequency division duplex (FDD) based systems, the DL and UL channels occupy different frequency bands [73]. In this case, estimation of the DL channel using UL training sequences is not possible. To obtain the CSI

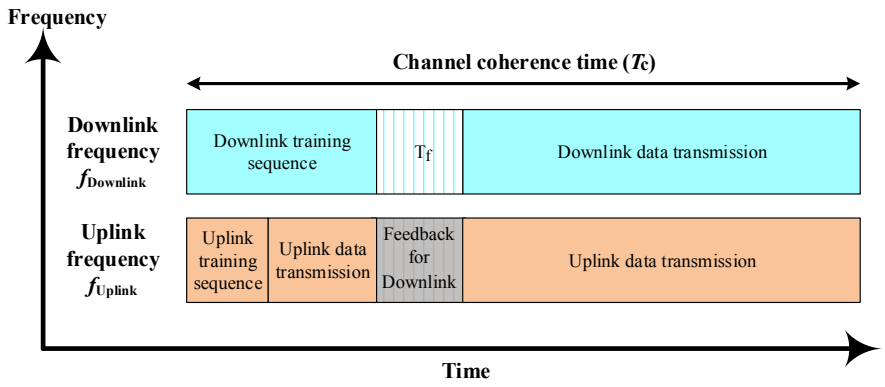


Figure 2.4 Separate downlink and uplink training and data payload phases of the block fading model in FDD operation mode.

at the BS in FDD based systems, the mobile users would need to estimate the DL channels of each of the N base station antennas and send the quantized channel estimates back to the BS to design the precoder. This is generally deemed unfeasible for the FDD massive MIMO systems with large N since the overhead for the DL channel estimation is proportional to the number of BS antennas N [27]. As such, the available coherence time would be largely occupied by the channel training, leaving insufficient time for transmitting useful data to the users [27, 74]. Therefore, there is an essential requirement in FDD massive MIMO systems to minimise the training sequence length for DL channel estimation over a limited coherence time. Fig. 2.4 illustrates the basic coherence block length with the slot structure for DL and UL channels in FDD operation mode. The parameters in the Fig. 2.4 are T_c, which is defined as the channel coherence time in symbols per each transmission block, T_f, which is defined as the feedback time and T_d, which is defined as the time duration used for transmitting useful data to users. In particular, the available channel coherence time is divided into the training duration, the feedback time and the data transmission duration. As such, the remaining time duration $T_d = T_c - T_p - T_f$ in symbols correspond to data transmission. As the purpose of this thesis is to concentrate on the DL training sequence design, minimising the training sequence length over a limited coherence time, the UL feedback time T_f and associated error rate are assumed to be zero. The amount of time for UL feedback is expected to be relatively small if an efficient feedback design is considered [75, 76], for example, a quantised codebook design.

2.4 System performance metric

The definition of the data traffic throughput, which corresponds to the maximum amount of data rate provided by the communication system, is given by

$$\text{Throughput (bit/s)} = \text{bandwidth (Hz)} \times \text{spectral efficiency (bit/s/Hz)}. \tag{2.2}$$

From (2.2), enhancing the system throughput requires either allocating more bandwidth or improving the spectral efficiency. The spectral efficiency, also known as the achievable sum rate or capacity, is defined as the number of information bits that can be reliably transmitted over the channel [39]. Throughout this thesis, the achievable sum rate performance metric is used for massive MIMO systems evaluation. Accordingly, a DL achievable sum rate, C, can be expressed as

$$C = \sum_{k=1}^{K} \log_2\big(1 + \text{SINR}_k\big), \tag{2.3}$$

where SINR_k denotes the associated received SINR term at the k-th UT. In information theory, the SINR is a quantity used to provide theoretical upper bounds on channel capacity (or the amount of rate of information transfer) in the wireless communication systems. Under the block fading channel model and imperfect channel estimation, a pre-log fraction $\big(1 - T_\text{p}/T_\text{c}\big)$ should be considered, accounting for the achievable sum rate loss due to the DL channel training. The ultimate goal of this thesis is to investigate how the DL achievable sum rate of an FDD massive MIMO system can be maximised with limited coherence time using a low-complexity training sequence design and minimum training duration.

2.4.1 Precoding techniques

Typically, the achievable sum rate is degraded by the interfering signals from other transmissions, which are simultaneously propagated over the same time and frequency. Therefore, spatial signal preprocessing/precoding at the BS is necessary to mitigate the interference and to form a fixed beam towards the desired user. Fig. 2.5 shows a massive MIMO system architecture with single stage digital precoding. The precoding scheme allows to adopt an antenna

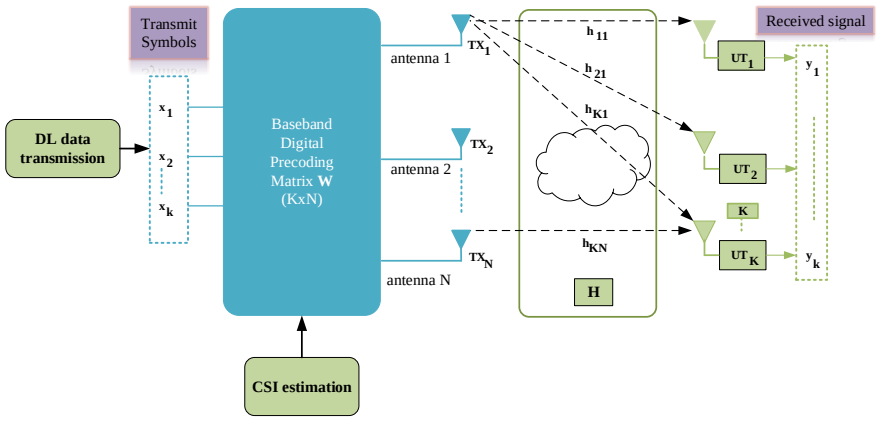


Figure 2.5 Schematic of massive MIMO architecture with single stage digital precoding at the BS.

array to transmit one or multiple spatially directive signals where the phases and powers can be jointly designed to achieve high achievable sum rate through spatial multiplexing of many UTs. The precoding technique plays an essential role in the spatial formatting of transmitted signals since it affects the channel characteristics in terms of directivity and gain of the desired and interfering signals [77]. The key requirement for efficient precoding design is to have accurate CSI at the BS, as discussed in subsection 2.2.2.

Information theory demonstrates that the optimum precoding strategy for multiuser MIMO systems can be achieved through the use of a theoretical pre-interference cancellation technique known as dirty-paper coding (DPC) [78–80]. However, DPC precoding is a non-linear technique that requires high computational complexity and to date is considered to be unfeasible for practical applications [32]. Alternatively, linear precoding techniques could be used to achieve a sub-optimum performance while reducing the computational complexity substantially [81–83]. Two types of linear precoders, the beamforming (BF) or maximum ratio transmitter (MRT) precoder and the regularised zero forcing (RZF) precoder [1] are considered throughout this thesis. Such linear precoding techniques are commonly prevailing and frequently encountered in the open literature on massive MIMO system evaluation. This has motivated the author to investigate the performance of the massive MIMO systems using such types of precoders. The associated SINR term at the k-th UT and the linear precoders will be provided in Chapter 3.

[1] The BF and RZF precoders are defined in Chapter 1 Section 1.7.

2.5 Mathematical tools

This section presents the essential mathematical tools that are used in this thesis. Matrix decomposition is discussed in subsection 2.5.1 while random matrix theory is briefly introduced in subsection 2.5.2. Further details on these tools can be found in [84, 85].

2.5.1 Matrix decomposition

A useful principle in matrix analysis is the decomposition, in which a matrix can be transformed into the product of orthogonal/unitary matrices and a diagonal form using eigenvalue decomposition (EVD) or singular value decomposition (SVD) [86]. This mathematical transformation plays an essential role in analysing the MIMO vector channels (matrices) and obtaining the degrees of freedom (DoF), which are intrinsic to these channels.

Suppose that $\mathbf{R}$ is an $N \times N$ complex Hermitian covariance matrix. By means of EVD, any $N \times N$ matrix can be factored as

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}}, \tag{2.4}$$

where $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_N] \in \mathbb{C}^{N \times N}$ is a complex unitary matrix whose columns are the eigenvectors of $\mathbf{R}$ and $\mathbf{\Lambda} \in \mathbb{R}^{N \times N}$ is a diagonal matrix whose diagonal entries are the eigenvalues of $\mathbf{R}$ arranged in descending order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$. The unitary matrix of the eigenvectors is orthogonal and satisfies $\mathbf{U}^{\mathrm{H}}\mathbf{U} = \mathbf{I}_N$. Since matrix $\mathbf{R}$ is a Hermitian matrix, the entries on the main diagonal of $\mathbf{\Lambda}$ are non-negative [84]. For a rank deficient matrix, a rank of $r = \mathrm{rank}(\mathbf{R})$ with $r \leq N$ equals to the number of non-zero eigenvalues, which corresponds to the strongest eigenmodes (eigendirections) of $\mathbf{R}$. The first r columns space eigenvectors of matrix $\mathbf{R}$, i.e., $\mathbf{U} \in \mathbb{C}^{N \times r}$, is known as the span of $\mathbf{R}$, which is associated to the non-zero eigenvalues of $\mathbf{R}$. The null space of a matrix $\mathbf{R}$ is the last $N - r$ dimension subspace of the eigenvalues of $\mathbf{R}$, which corresponds to last eigenvector columns of $\mathbf{R}$ [84]. The trace operator of a square matrix $\mathbf{R}$, denoted by $\mathrm{tr}(\mathbf{R})$, equals to the sum of diagonal entries (eigenvalues) of $\mathbf{R}$. The trace is a linear mathematical operation, which is invariant to unitary rotation and a change of basis. The eigendecomposition of the channel covariance matrix is a useful tool that will be utilised in this

thesis for designing the training sequences of massive MIMO systems.

2.5.2 Random matrix theory

In wireless communication systems, the channel responses depend on the propagation environments, which in general vary randomly over the time and frequency. Hence, random distribution is typically used in modelling the channel variations [39]. Typically, modelling the channel variations relies on computationally demanding Monte Carlo simulations, thereby obtaining precise engineering insight and explicit tight bounds as the evaluation of the system performance can become an intractable problem. A computationally feasible solution for characterising the random channel distribution can be obtained by invoking a large system limit based on the well established field of random matrix theory. The large system limit based on random matrix theory methods has been widely used in the performance analysis of wireless communication systems [85]. In the random matrix theory, an asymptotic approximation can be obtained when two of the system parameters, e.g., number of BS antennas and number of users, go to infinity at a finite constant ratio. This is because the eigenvalue distribution of the large random matrices converge asymptotically to a fixed distribution. This essential property of the large system limit based on random matrix theory methods can be utilised to provide an analytical closed-form expression for the system performance. While the resulting analytical expressions are only tight asymptotically, they usually provide an accurate approximation for the exact simulation values with a finite number of system parameters.

Throughout this thesis, a random matrix theory method is used to avoid executing extensive Monte Carlo simulations, thus obtaining analytical solutions for the optimisation problems considered with low-complexity. A random matrix theory method allows a mathematical analysis simplification to be achieved for advanced channel models, such as those encountered in arbitrary correlation matrices. The system performance under these types of channel models is challenging to analyse but becomes tractable with the random matrix theory [32]. To this end, an asymptotically tight approximation of the SINR_k at the k-th UT, denoted $\overline{\mathrm{SINR}}_k$ can be obtained as indicated in (2.5) when N and K grow without bound while the ratio $K/N > 0$ is kept constant.

$$\mathrm{SINR}_k - \overline{\mathrm{SINR}}_k \xrightarrow[N\to\infty]{} 0 \tag{2.5}$$

Consequently, an explicit closed-form expression for the DL achievable sum rate of a massive MIMO system can be provided for both the BF and RZF precoders considered.

2.6 Rayleigh fading channel

As discussed earlier, a random distribution is used for modelling the channel variations in the wireless communication systems. The channel statistics that correspond to the distributions of the random variables are typically assumed to be available everywhere in the network [39]. This can be justified by the fact that it is sufficient to know the full distributions of random variables, which correspond to their first- and second-order moments [87]. Typically, the first-order moment of the channel statistic denotes the channel mean while the second-order moment refers to the channel covariance matrix [39].

A well-known approach for modelling the channel covariance matrix is the uncorrelated Rayleigh fading model in which the channel coefficients are assumed to be independent and uniformly distributed over the unit-sphere [28]. In probability theory and central limit theorem, the linear superposition of many independent random variables allow the channel coefficients to be approximated by a circular symmetric complex Gaussian distribution. The Rayleigh fading channel arises in scenarios of rich scattering geometry, in which the signal can be propagated through a large number of independent scattering objects. In the Rayleigh fading channel, there should be a significant number of multipath components in the angular domain, and the energy should be equally spread out in all directions [73]. This is in contrast to the correlated channel, discussed in detail in the next section, where there are only a few paths in some of the angular directions. The condition for the channel characteristics to be spatially uncorrelated as exhibited in rich scattered environment seems very strict. Further, due to the use of a large number of BS antennas, which are closely spaced by half a wavelength, the channel coefficients in massive MIMO systems may be correlated. Therefore, to provide a realistic performance assessment of massive MIMO systems, channel models that reflect the key characteristics of a large number

of antenna elements are desired. The following section reviews spatial correlation in channel models.

2.7 Spatial channel correlation

This section reviews the spatial correlation in MIMO channels. The BS antennas might have non-uniform radiation patterns and varying polarization with non-isotropic transmission, making the MIMO channels spatially correlated [49, 88–90]. This is in contrast to the spatially uncorrelated channels, discussed in Section 2.6, where each single antenna which has its own pattern, radiates isotropically over the unit-sphere. In practice, field measurements have shown that the MIMO channel coefficients are spatially correlated in outdoor and indoor propagation environments [49, 88–90]. When the amount of scattering around the UT in the physical propagation environment is limited, the signals can be propagated through a few strong spatial directions [91]. As such, the channel can be partitioned into a finite number of dimensions in the angular domain, which is relatively small compared to the number of BS antennas. The signals from the transmitter to the receiver would only occupy these spatial directions. The energy would also be concentrated into these strong spatial directions and not isotropically distributed in all directions. The spatial correlation is beneficial to the channel estimation [92]. This is because the correlations would increase the spatial resolution and reduce the uncertainty in channel dimensions that the signals propagate through, thus, improving the CSI accuracy.

In this thesis, two commonly used spatial channel correlation models are considered: the analytical P-degrees of freedom (P-DoF) and the scattering one ring (OR) channel models. These channel models are frequently encountered in the open literature on MIMO evaluation [30, 93–96], and thus, will be used in the massive MIMO evaluation here. These channel models incorporate stochastic rank deficiency. Channels with rank deficient covariance matrices were called stochastically rank deficient in [97] and the same terminology is used in this thesis.

2.7.1 The analytical P-DoF channel model

The correlation model depends on the number of degrees of freedom offered by the physical channel, which can be much smaller than the number of BS antennas N. Therefore, the correlation model can be modelled as a P-dimensional subspace, where P is the number of angular spatial directions or bins. These angular bins correspond to the number of significant multipaths in the angular domain. In this thesis, an analytical physical channel model is considered, where the angular domain is separated into a finite P number of directions, i.e. P-DoF. The P-DoF channel model is used to obtain a closed-form solution for the achievable sum rate in the BF and RZF precoders. This model also allows the optimum training sequence length for the BF precoder to be mathematically derived and expressed in an analytical form. The P-DoF channel model for the system under consideration is written as [30]

$$\mathbf{h}_k = \sqrt{\frac{N}{P}}\mathbf{A}_k\mathbf{b}_k, \tag{2.6}$$

where $P/N \in (0, 1]$, the elements of $\mathbf{b}_k \in \mathbb{C}^P$ are i.i.d. $\mathcal{CN}(0, 1)$ and $\mathbf{A}_k \in \mathbb{C}^{N\times P}$ is constructed from $P \leq N$ columns of an arbitrary $N \times N$ unitary matrix so that $\mathbf{A}_k^{\mathrm{H}}\mathbf{A}_k = \mathbf{I}_P$. Clearly, $\mathbf{R}_k$ has rank P and the channel is stochastically rank deficient if $P < N$. In general, the ratio P/N controls the degrees of freedom of the channel and represents the extent of correlation or amount of scattering in the channel, and thus, models the radio environment. The smaller P/N is, the more concentrated the channel energy is in the non-zero eigenmodes and the more correlated the channel is. The normalization factor in (2.6) guarantees that the total average power of the channel is $\mathbb{E}\big[\mathrm{tr}(\mathbf{H}^{\mathrm{H}}\mathbf{H})\big] = \sum_{k=1}^{K}\mathrm{tr}\big(\mathbb{E}[\mathbf{h}_k\mathbf{h}_k^{\mathrm{H}}]\big) = \sum_{k=1}^{K}\mathrm{tr}(\mathbf{R}_k) = NK$. The covariance matrices of the users for the channel defined in (2.6) are of the form $\mathbf{R}_k = \mathbb{E}[\mathbf{h}_k\mathbf{h}_k^{\mathrm{H}}] = \frac{N}{P}\mathbf{A}_k\mathbf{A}_k^{\mathrm{H}}$, where the expectation is taken over $\mathbf{b}_k$. As an expectation of the instantaneous channel vectors, which depends on the nature of the scattering environment and the propagation geometry, the channel covariance matrix is considered to be locally stationary, varying more slowly than the instantaneous channel of the coherence time [39]. Therefore, the channel covariance matrix may be accurately estimated by either the FDD or TDD schemes considered [98–105].

While the channel of different users associated with the BS can be random, they might exist in mutually orthogonal subspaces. This behaviour of the channel is known as a non-overlapping

angle of arrival (AoA) support [52, 106–108], in which the AoA of desired and interfering users are disjoint. In practice, the scenario of the non-overlapping signal subspaces between the desired and the interfering users can be realised when the ULA at the BS has a large number of antenna elements and when the users are well separated in the angular domain [106]. In the multiuser scenario with non overlapping AoA supports, the covariance matrices of the users are all statistically different and independent of each other.

Recalling the EVD of the channel covariance matrix is given in (2.4) and the definition of non-overlapping AoA supports [106], in which the channel covariance matrices can be asymptotically orthogonal and satisfy

$$\mathbf{U}_k^{\mathrm{H}} \mathbf{U}_l = 0, \ \forall l \neq k, \qquad \text{as} \ \ N \rightarrow \infty. \tag{2.7}$$

The unitary symmetric structure of the discrete Fourier transform (DFT) matrix can be used to obtain perfectly non-overlapping AoA supports with the P-DoF channel model (2.6). In contrast, when the users exhibit completely overlapping AoA supports, the covariance matrices of the users are all statistically the same and of the form $\mathbf{R}_k = \mathbf{R} \ \forall k = 1, \ldots, K$.

2.7.2 The OR scattering channel model

Section 2.7.1 presented an analytical P-DoF channel model, which facilitates analytical investigations. For practical purposes, this subsection considers the application of an alternative channel model called the OR scattering channel model, which is a more realistic channel model [94]. The OR scattering channel model represents an environment where all scatterers are located on a ring around the UT and there is no local scattering around the BS. The system geometric parameters of the channel covariance matrix $\mathbf{R}_k$ in the OR scattering channel model are determined by the angular spread in the azimuth direction ω_{H}, the azimuth angles of arrival θ_k, and normalised antenna spacing D in wavelengths. Specifically, the (m,n)th element of $\mathbf{R}_k$ is given in Toeplitz form [94] as

$$[\mathbf{R}_k]_{m,n} = \frac{1}{2\omega_{\mathrm{H}}} \int_{-\omega_{\mathrm{H}}+\theta_k}^{\omega_{\mathrm{H}}+\theta_k} e^{-j2\pi \, D \, (m-n) \sin(\Delta)} d\Delta. \tag{2.8}$$

where Δ represents the intervals/ranges of the AoAs distribution (i.e. $\Delta \in [-\pi/3, \pi/3]$) since a 120° sector is considered in this thesis. The integration in (2.8) is computed numerically. Due to the different scattering geometries associated with each user's position in a geographic area, the angular support of each user's channel appears random. The randomness in the user locations captures the fact that the angular supports of the users may partially overlap. In addition, when the BS antennas are closely spaced and the amount of scattering around the UT is limited, as indicated by both D and ω_{H} being small, some of the eigenvalues of $\mathbf{R}_k$ are close to zero, making $\mathbf{R}_k$ effectively rank deficient. In contrast to the P-DoF channel model, described in Section 2.7.1, in the OR model the non-zero eigenvalues are not necessary all equal.

An approximation of the actual rank $r_k \in \mathbb{Z}^+$ for large but finite N that contain the effective non-zero eigenvalues of the channel covariance matrix in (2.8) is given by [51]

$$r_k = N\beta_k, \tag{2.9}$$

where $\beta_k = \min\{1, f(D, \omega_{\mathrm{H}}, \theta_k)\}$ is the asymptotic normalised rank of the channel covariance matrix $\mathbf{R}_k$ with $f(D, \omega_{\mathrm{H}}, \theta_k) = \lfloor D\sin(\theta_k - \omega_{\mathrm{H}}) - D\sin(\theta_k + \omega_{\mathrm{H}})\rfloor$. While (2.9) can accurately predict the rank of the OR channel model that is related to the number of non-zero eigenvalues of $\mathbf{R}_k$, this number may differ between users due to different θ_k. The DL training sequence design and channel estimation in the massive MIMO systems is still an ongoing area of research, which involves many theoretical and practical challenges.

2.8 Summary

This chapter has introduced the necessary background materials that are relevant to the thesis. This chapter explained that accurate CSI estimation is essential for efficient precoding design. Obtaining the CSI at the BS depends on the duplexing mode of the system, i.e. either FDD or TDD mode. Despite the promising results of TDD operation with massive MIMO when it comes to the CSI acquisition, the vast majority of the currently deployed cellular networks operate in FDD mode. However, in massive MIMO with a large number of BS antenna elements, the training overhead required for obtaining the CSI is overwhelming in FDD based systems.

Therefore, enabling FDD mode, which still dominates current cellular networks, to operate in massive MIMO systems is essential.

This chapter has demonstrated that when the BS antenna elements are closely spaced and the mobile users experience limited scattering geometries, the channels could be strongly correlated. Increasing the level of correlation in the channels results in a rank deficient covariance matrix. Exploiting the spatial channel correlation in the CSI estimation is essential for designing feasible FDD massive MIMO systems. This is because the rank of the transmit correlation matrix can be scaled sublinearly with N, resulting in feasible pilot overhead requirements for FDD operation with massive MIMO systems. Specifically, in the presence of spatially correlated fading; it is possible to train the channel for a minimum duration upper bounded by the rank of the corresponding transmit spatial correlation matrix, to achieve a decrease in the CSI estimation error with decreasing noise power. This chapter has also explained that users may have distinct spatial correlation patterns, which could arise due to independent propagation conditions and scattering geometries between users. However, optimising the training sequences for CSI estimation in a general scenario with heterogeneous user channels is more challenging than in the scenario with common user channel correlation.

In the next chapter, a gradient-based iterative algorithm for DL training sequence optimisation will be investigated with an aim to reduce the computational complexity.

Chapter 3

Gradient-Based Iterative Algorithms for FDD Training Sequence Design

3.1 Introduction

Chapter 2 indicated that in non-reciprocal channels, such as those encountered in FDD based systems, CSI estimation using DL training sequences is very challenging. It also explained that in order to obtain sufficiently accurate CSI estimation in the scenario where the users exhibit different spatial correlation patterns, i.e., span distinct channel covariance matrices, the DL training sequences should be optimised at the BS. One approach to optimising the training sequences at the BS with single-stage precoding and heterogeneous user channels is to use gradient-based iterative algorithms. For example, in [53] and [56] the DL training sequences are optimised iteratively as a solution to a SCMI maximisation problem and a SMSE minimisation problem, respectively. Specifically, in [53], a gradient-based iterative algorithm based on the Lagrange multiplier on the Euclidean space is proposed to optimise the DL training sequence that maximises the SCMI between the channel and the channel estimates. In [56], an iterative steepest descent algorithm based on the complex Grassmannian space is developed to optimise the DL training sequence that minimises the SMSE design criterion. While the investigations in [53] and [56] have optimised the DL training sequences iteratively in order to maximise the SCMI and minimise the SMSE, respectively, the algorithms used in the optimi-

sation exhibit slow convergence speed. However, in real-time systems, the CSI estimation must be obtained more frequently with an acceptable latency, especially when the channel exhibits a shorter coherence time. Therefore, there is an essential requirement for developing a new iterative algorithm that optimises the DL training sequence, i.e., maximises the achievable sum rate, while at the same time exhibits fast convergence speed so that to obtain the CSI estimation more quickly.

In wireless communication systems and array signal processing, several studies have considered various optimisation problems under a unitary matrix constraint [109–112]. Consequently, advanced gradient-based iterative algorithms have been developed to find efficient solutions with fast convergence rate to such optimisation problems [57, 113–115]. Some of those efficient solutions have been obtained using a geometrical approach based on the smooth parameter space known as the Riemannian manifold [116–120]. The potential of the Riemannian manifold has motivated the author of this thesis to explore the application of such a geometrical approach for optimising the DL training sequences in an FDD massive MIMO system when the users exhibit distinct spatial channel correlations.

3.1.1 Contributions and chapter findings

This chapter proposes a geometrical spaced conjugate gradient descent (CGD) optimisation algorithm, which uses the Riemannian manifold for optimising the DL training sequences in order to speed up the convergence rate for fast CSI estimation in the FDD massive MIMO systems and hence reduce the overall precoder complexity. In addition, this chapter develops the computational complexity analyses of the proposed CGD algorithm and the state-of-the-art SCMI [53] and SMSE [56] gradient-based iterative algorithms. Comparisons are presented for the overall complexity and sum rate performances between the proposed CGD algorithm and existing iterative algorithms for DL sequence optimisation in both BF and the RZF precoders. For a fair comparison, the same error tolerance for stopping criteria is used for all the training sequence designs. The complexity analyses are obtained based on the number of multiplications and additions in flops [121–123] per iteration for the three training sequence designs considered. Further, this chapter designs and model an FDD transmission based on massive MIMO transmission system and developing a physical layer model using the MATLAB software sim-

ulator. To this end, a massive MIMO system model that characterises the achievable sum rate performance based on BF and RZF is established. To demonstrate the impact of spatial correlation on the DL CSI estimation of an FDD massive MIMO system, and hence the sum rate, the scattering one ring channel model is used. The results show that the proposed CGD algorithm improves the convergence behaviour of the existing gradient-based iterative algorithms due to the use of matrix exponential search on the Riemannian manifold. This improvement results in an optimised DL training sequence for the CSI estimation in an FDD multiuser massive MIMO system while improving the convergence speed and hence reducing the overall computational complexity. Importantly, this chapter is essential to provide a useful indication of the best sum rates that can be achieved in the FDD massive MIMO systems, which more practical solutions, as developed in this thesis, can aim for.

3.2 System model description

The present model in this chapter considers a single-cell DL mobile wireless communications system, where the BS is equipped with a uniform linear array (ULA) of N transmit antennas that serves K single antenna UTs with $N \gg K$. Due to the single antenna UTs, no correlation occurs at the receiver. Without loss of generality, non-line-of-sight (NLOS) Rayleigh fading channels over a single-frequency band are considered with an overall coherence time denoted by $T_{\mathrm{c}} \in \mathbb{Z}^{+}$ and enumerated in symbols per transmission block. Each coherence block contains a period of time for data transmission, the DL received signal during the data transmission phase at the k-th UT is given by [29, 30]

$$y_k = \sqrt{\rho_{\mathrm{d}}}\mathbf{h}_k^{\mathrm{H}}\mathbf{s} + n_k, \tag{3.1}$$

where $\mathbf{h}_k \in \mathbb{C}^N$ is the DL complex channel vector intended between the BS and the k-th UT. On the other hand, additive noise n_k is modelled as an independent and identically distributed zero mean unit variance circularly symmetric complex Gaussian (CSCG) random variable. The

DL transmit vector $\mathbf{s} \in \mathbb{C}^N$ from the BS can be written as

$$\mathbf{s} = \sqrt{\xi}\mathbf{W}\mathbf{x}, \tag{3.2}$$

where $\mathbf{W} = \left[\mathbf{w}_1, \ldots, \mathbf{w}_K\right] \in \mathbb{C}^{N\times K}$ is the complex precoding matrix at the transmit BS and $\mathbf{x} = \left[x_1, \ldots, x_K\right]^{\mathrm{T}} \in \mathbb{C}^K$ is independently and identically distributed with a zero mean CSCG vector of data symbols, which satisfies $\mathbb{E}\left[\mathbf{x}\mathbf{x}^{\mathrm{H}}\right] = \mathbf{I}_K$. When the matrix power normalisation technique is used [30, 124], the normalisation constant parameter ξ can be written as

$$\xi = \frac{K}{\mathbb{E}\left[\mathrm{tr}\left(\mathbf{W}\mathbf{W}^{\mathrm{H}}\right)\right]}, \tag{3.3}$$

which satisfies that $\mathbb{E}\left[\|\mathbf{s}\|^2\right] = K$ and the average BS transmit power allocated per UT is ρ_{d} during the data transmission phase. In massive MIMO systems, the UT in general does not know the exact channel vector and beamforming vector [29, 30], but instead estimates their average effect through $\sqrt{\xi}\mathbb{E}\left[\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k\right]$ as considered in [29, 30]. This is obtained by determining the mean of an inner products of the effective channel for the k-th UT and the beamforming vector, which is widely used in the massive MIMO literature and becomes a part of the received DL effective SINR calculation. The DL received signal at the k-th UT in (3.1) can be decomposed as

$$\begin{aligned} y_k = \sqrt{\rho_{\mathrm{d}}\,\xi}\,\mathbb{E}[\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k]x_k + \sqrt{\rho_{\mathrm{d}}\,\xi}\left(\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k - \mathbb{E}[\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k]\right)x_k \\ + \sqrt{\rho_{\mathrm{d}}\,\xi}\,\mathbf{h}_k^{\mathrm{H}}\sum_{l\neq k}^{K}\mathbf{w}_l x_l + n_k, \end{aligned} \tag{3.4}$$

where n_k is the additive independent received noise, as defined above. The first term in (3.4) denotes the information signal of interest at the k-th UT (i.e. the desired signal), which is received as an average over the precoded channel. This is considered to be a part of the received effective SINR calculation. The second term represents the error caused by the imperfect channel knowledge at the UT (i.e. treated as an uncorrelated noise signal). The third term is the residual multiuser interference after precoding.

From (3.4), the received effective SINR term of the fading channel from the transmit BS to the k-th UT with linear precoding is given as

$$\mathrm{SINR}_k = \frac{\xi \mid \mathbb{E}[\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k] \mid^2}{\frac{1}{\rho_{\mathrm{d}}} + \xi\mathbb{E}\left[\mid \mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k - \mathbb{E}[\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k] \mid^2\right] + \xi \sum_{l\neq k}^{K} \mathbb{E}\left[\mid \mathbf{h}_k^{\mathrm{H}}\mathbf{w}_l \mid^2\right]}, \tag{3.5}$$

where $|\cdot|^2$ represents the squared absolute of a complex number, (i.e., the power of a signal), the term $1/\rho_{\mathrm{d}}$ denotes the inverse of the per-user SNR during the data transmission phase. The expectation (i.e., the averaging process) in (3.5) is taken with respect to different channel realisations, which are obtained separately by means of Monte Carlo simulation.

Remark 3.1. Although the performance evaluation of an FDD massive MIMO system under consideration in this chapter is carried out based on Monte Carlo simulations, a large system limit based on random matrix theory methods will be adopted later in Chapter 5 to avoid executing such extensive simulations. In particular, random matrix theory methods will be used in Chapter 5 to provide an asymptotically tight approximation of the SINR expression in (3.5).

In general, the received effective SINR term in (3.5) depends on the channel statistics, the channel estimates, and the linear precoding technique used at the BS. Based on the discussion in Chapter 2, two linear precoders, the BF or MRT precoder and the RZF precoder are considered as defined in (3.6) [30].

$$\mathbf{W} = \begin{cases} \hat{\mathbf{H}}^{\mathrm{H}} & \text{for BF/MRT,} \\ \hat{\mathbf{H}}^{\mathrm{H}}\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathrm{H}} + N\zeta\mathbf{I}_K\right)^{-1} & \text{for RZF,} \end{cases} \tag{3.6}$$

where ζ is the regularisation parameter for the RZF precoder, which is considered to be the inverse of the per-user SNR, $1/\rho_{\mathrm{d}}$ [30] and N denotes the number of BS antennas. The term $\hat{\mathbf{H}}$ is the estimate, discussed in detail in the next section, of the DL channel $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_K]^{\mathrm{H}} \in \mathbb{C}^{K\times N}$. The channel vectors $\mathbf{h}_k$, $k = 1, \ldots, K$, are considered to be independently and identically randomly distributed CSCG with zero mean and the k-th UT's correlation matrix, $\mathbf{R}_k$, at the transmit BS-side is given by $\mathbf{R}_k = \mathbb{E}\left[\mathbf{h}_k\mathbf{h}_k^{\mathrm{H}}\right] \in \mathbb{C}^{N\times N}$. The total average power of the channel is normalised as $\mathbb{E}\left[\mathrm{tr}(\mathbf{H}^{\mathrm{H}}\mathbf{H})\right] = \sum_{k=1}^{K} \mathrm{tr}\left(\mathbb{E}[\mathbf{h}_k\mathbf{h}_k^{\mathrm{H}}]\right) = \sum_{k=1}^{K} \mathrm{tr}(\mathbf{R}_k) = NK$. The following section explains the CSI estimation process based on DL training sequences together with problem formulation.

3.3 Channel estimation using DL training sequences in FDD

From (3.5) and (3.6) the sum rate performance metric in (2.3), described in Chapter 2 Section 2.4, depends on both channel statistics and channel estimation through the associated SINR term in (3.5). To gain further insight to the problem of DL CSI estimation, this section investigates the training sequence design in an FDD multiuser massive MIMO system. This is in contrast to UL channel estimation considered conventionally for reciprocal channels, such as those encountered in the appropriately calibrated TDD massive MIMO systems [30].

To estimate the DL channel, the BS transmits predetermined common pilot matix of duration or length T_p during the training-phase. To this end, the received signal during training-phase, $\mathbf{y}_k \in \mathbb{C}^{T_\mathrm{p}}$, at the k-th user is given by [56]

$$\mathbf{y}_k = \sqrt{\rho_\mathrm{p}}\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{h}_k + \mathbf{n}_k, \tag{3.7}$$

where $\mathbf{S}_\mathrm{p} \in \mathbb{C}^{N\times T_\mathrm{p}}$ is the spatio-temporal common pilot matrix, which is normalised as $\mathrm{tr}\big(\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{S}_\mathrm{p}\big) = T_\mathrm{p}$ so that the average transmitted power during the training-phase is equal to ρ_p. On the other hand, the receiver noise $\mathbf{n}_k \in \mathbb{C}^{T_\mathrm{p}}$ exhibits a CSCG distribution $\mathcal{CN}\big(\mathbf{0}, \mathbf{I}_{T_\mathrm{p}}\big)$. Since the channel vector $\mathbf{h}_k$ follows a CSCG distribution with known statistics at the BS, the MMSE estimator is used [25], and the resulting channel estimate $\hat{\mathbf{h}}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Phi}_k)$ is obtained as

$$\hat{\mathbf{h}}_k = \sqrt{\rho_\mathrm{p}}\mathbf{R}_k\mathbf{S}_\mathrm{p}\big(\rho_\mathrm{p}\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{R}_k\mathbf{S}_\mathrm{p} + \mathbf{I}_{T_\mathrm{p}}\big)^{-1}\mathbf{y}_k. \tag{3.8}$$

Using the MMSE estimator allows the k-th user's channel vector to be deconstructed into the channel estimate $\hat{\mathbf{h}}_k$ and the channel estimation error $\tilde{\mathbf{h}}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_k - \mathbf{\Phi}_k)$ as shown in (3.9).

$$\mathbf{h}_k = \hat{\mathbf{h}}_k + \tilde{\mathbf{h}}_k \tag{3.9}$$

The vectors $\hat{\mathbf{h}}_k$ and $\tilde{\mathbf{h}}_k$ are uncorrelated and $\mathbf{\Phi}_k \in \mathbb{C}^{N\times N}$ denotes the k-th user's covariance matrix with MMSE channel estimate $\forall k$ as expressed by [25]

$$\mathbf{\Phi}_k = \sqrt{\rho_\mathrm{p}}\mathbf{R}_k\mathbf{S}_\mathrm{p}\big(\rho_\mathrm{p}\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{R}_k\mathbf{S}_\mathrm{p} + \mathbf{I}_{T_\mathrm{p}}\big)^{-1}\sqrt{\rho_\mathrm{p}}\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{R}_k. \tag{3.10}$$

The MMSE channel estimation $\mathbf{\Phi}_k$ in (3.10) can be rewritten as in (3.11)

$$\mathbf{\Phi}_k = \mathbf{R}_k \mathbf{S}_\mathrm{p} \left(\mathbf{S}_\mathrm{p}^\mathrm{H} \mathbf{R}_k \mathbf{S}_\mathrm{p} + \frac{1}{\rho_\mathrm{p}} \mathbf{I}_{T_\mathrm{p}} \right)^{-1} \mathbf{S}_\mathrm{p}^\mathrm{H} \mathbf{R}_k. \tag{3.11}$$

The expression in (3.11) characterises the output of the channel estimator that minimises the MSE of the channel estimate. The overall MSE performance depends on the structure of the training sequence matrix $\mathbf{S}_\mathrm{p}$, training duration T_p and the training power ρ_p. This chapter proposes a low-complexity conjugate gradient descent iterative algorithm, described in detail in Section 3.4, to optimise $\mathbf{S}_\mathrm{p}$ for pre-selected training length T_p, and training power ρ_p. The proposed algorithm aims to achieve a robust sum rate performance when the users exhibit different correlations while at the same time reduces the computational complexity. Let $f(\mathbf{S}_\mathrm{p}) \in \mathbb{R}$ denote the sum MSE of the channel estimate cost function for all the users, which needs to be minimised with respect to the DL training matrix $\mathbf{S}_\mathrm{p}$, as given in [21, 56]

$$f(\mathbf{S}_\mathrm{p}) = \sum_{k=1}^{K} \mathrm{tr}\big(\mathbf{R}_k - \mathbf{\Phi}_k\big). \tag{3.12}$$

Substituting the covariance matrix with MMSE channel estimation $\mathbf{\Phi}_k$, which is given by (3.11), into (3.12), and using the trace operation property [125][1], the cost function in (3.12) can be reformulated as

$$f(\mathbf{S}_\mathrm{p}) = \sum_{k=1}^{K} \mathrm{tr}\big(\mathbf{R}_k\big) - \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{R}_k \mathbf{S}_\mathrm{p} \left(\mathbf{S}_\mathrm{p}^\mathrm{H} \mathbf{R}_k \mathbf{S}_\mathrm{p} + \frac{1}{\rho_\mathrm{p}} \mathbf{I}_{T_\mathrm{p}} \right)^{-1} \mathbf{S}_\mathrm{p}^\mathrm{H} \mathbf{R}_k \right). \tag{3.13}$$

3.3.1 Formulation of the channel estimation optimisation problem

Minimising the sum MSE cost function over the training sequence $\mathbf{S}_\mathrm{p}$ with a unitary matrix constraint for a given training-phase duration T_p in the DL multiuser massive MIMO system equates to the optimisation problem defined in (3.14).

$$\begin{aligned} \underset{\mathbf{S}_\mathrm{p}}{\text{minimise}} \quad & \sum_{k=1}^{K} \mathrm{tr}\big(\mathbf{R}_k\big) - \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{R}_k \mathbf{S}_\mathrm{p} \left(\mathbf{S}_\mathrm{p}^\mathrm{H} \mathbf{R}_k \mathbf{S}_\mathrm{p} + \frac{1}{\rho_\mathrm{p}} \mathbf{I}_{T_\mathrm{p}} \right)^{-1} \mathbf{S}_\mathrm{p}^\mathrm{H} \mathbf{R}_k \right) \\ \text{subject to} \quad & \mathbf{S}_\mathrm{p}^\mathrm{H} \mathbf{S}_\mathrm{p} = \mathbf{I}_{T_\mathrm{P}} \end{aligned} \tag{3.14}$$

[1] $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$

The sum MSE cost function, which corresponds to a function of the subspace that is spanned by the pilot matrix $\mathbf{S}_\mathrm{p}$, is invariant to the unitary rotation [56][2]. Consequently, the rotational invariant of the cost function can be expressed as in (3.15)

$$f(\mathbf{S}_\mathrm{p}\mathbf{U}) = f(\mathbf{S}_\mathrm{p}), \tag{3.15}$$

where $\mathbf{U}$ is any unitary matrix. This implies that the sum MSE cost function is a function of range of the pilot matrix $\mathbf{S}_\mathrm{p}$, so that the domain of the sum MSE cost function becomes a codomain of $\mathbf{S}_\mathrm{p}$, as explained in [126]. This property allows the Riemannian manifold approach to be used to solve the optimisation problem in (3.14). Due to (3.15), the unitary constraint in (3.14) is automatically met with the proposed CGD algorithm, and thus, the obtained training sequence has orthonormal columns (i.e., $\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{S}_\mathrm{p} = \mathbf{I}_{T_\mathrm{p}}$).

To apply the CGD optimisation algorithm on a Riemannian manifold, the partial derivative of the sum MSE cost function under consideration needs to be determined with respect to the DL training sequence matrix $\mathbf{S}_\mathrm{p}$. To this end, the partial derivative $\mathbf{\Gamma} \in \mathbb{C}^{N\times T_\mathrm{p}}$ of the sum MSE cost function $f(\mathbf{S}_\mathrm{p})$, defined in (3.13), with respect to $\mathbf{S}_\mathrm{p} \in \mathbb{C}^{N\times T_\mathrm{p}}$ is given by [56, 127]

$$\begin{aligned}\mathbf{\Gamma} = 2\sum_{k=1}^{K}\Bigg(&-\mathbf{R}_k^2\mathbf{S}_\mathrm{p}\Big(\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{R}_k\mathbf{S}_\mathrm{p} + \frac{1}{\rho_\mathrm{p}}\mathbf{I}_{T_\mathrm{p}}\Big)^{-1} + \mathbf{R}_k\mathbf{S}_\mathrm{p}\times \\ &\Big(\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{R}_k\mathbf{S}_\mathrm{p} + \frac{1}{\rho_\mathrm{p}}\mathbf{I}_{T_\mathrm{p}}\Big)^{-1}\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{R}_k^2\mathbf{S}_\mathrm{p}\Big(\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{R}_k\mathbf{S}_\mathrm{p} + \frac{1}{\rho_\mathrm{p}}\mathbf{I}_{T_\mathrm{p}}\Big)^{-1}\Bigg).\end{aligned} \tag{3.16}$$

where the right hand side in (3.16) is obtained by deriving the cost function in (3.14) with respect to the $\mathbf{S}_\mathrm{p}$, (i.e., $\partial f(\mathbf{S}_\mathrm{p})/\partial\mathbf{S}_\mathrm{p}^*$). The following section discusses in detail the proposed CGD iterative optimisation algorithm based on the Riemannian manifold.

[2]A random matrix $\mathbf{S}$ is unitary invariant if its distribution remains unchanged when multiplied by any unitary matrix that is independent of $\mathbf{S}$, i.e. $\mathbf{S} = \mathbf{S}\mathbf{U}$ [84].

3.4 Training sequence optimisation based on conjugate gradient descent algorithm on the Riemannian manifold

In this section, the conjugate gradient descent algorithm based on the Riemannian manifold is used to optimise the DL pilot matrix $\mathbf{S}_{\mathrm{p}}$ iteratively across multiple users with independent channel covariance matrices. In optimising (3.14) for $\mathbf{S}_{\mathrm{p}}$, the conjugate gradient descent method uses the almost periodic property of the matrix exponential search [119, 120, 128, 129] on the Riemannian manifold in order to increase the algorithm convergence speed by lessening the required number of iterations. In particular, the introduction of matrix exponential on the Riemannian manifold aids the proposed conjugate gradient descent algorithm to operate over a smaller dimensional search spaces, and thus, provides a fast convergence behaviour. Further details on the matrix exponential property and how it is locally parameterised on the Riemannian manifold can be found in [119, 120].

Algorithm 1 summarises the proposed CGD iterative algorithm, which optimises the DL training sequence $\mathbf{S}_{\mathrm{p}}$ in an FDD massive MIMO system for a given training-phase duration. The following steps explain Algorithm 1 in further details.

Algorithm 1 Iterative conjugate gradient descent (CGD)

1: Initialization: $t = 0$ and $\mu = 1, \mathbf{S}_t = \mathbf{I}$
2: Determine the gradient descent direction using (3.17), and set $\mathbf{G}_t = \mathbf{D}_t$.
3: If $\langle \mathbf{D}_t, \mathbf{D}_t \rangle < \epsilon$, then break.
4: Determine the matrix exponential on the Riemannian manifold: $\mathbf{P}_t = \exp(-\mu \mathbf{D}_t)$.
5: While $f(\mathbf{S}_t) - f(\mathbf{P}_t \mathbf{S}_t) \geq \mu \langle \mathbf{D}_t, \mathbf{D}_t \rangle$, then double $\mathbf{P}_t$, $\mu := 2\mu$.
6: While $f(\mathbf{S}_t) - f(\mathbf{P}_t \mathbf{S}_t) < 0.5\mu \langle \mathbf{D}_t, \mathbf{D}_t \rangle$, then halve $\mathbf{P}_t$, $\mu := 0.5\mu$.
7: Update $\mathbf{S}_{t+1} := \mathbf{P}_t \mathbf{S}_t$, and update $\mathbf{G}_{t+1} = \mathbf{D}_{t+1} + \upsilon_t \mathbf{G}_t$, $t := t + 1$, and go to step 2.

- Step 1 — Initial step: The proposed iterative gradient algorithm starts by selecting an initial training sequence matrix $\mathbf{S}_{\mathrm{p}} \in \mathbb{C}^{N \times T_{\mathrm{p}}}$, which satisfies $\mathbf{S}_{\mathrm{p}}^{\mathrm{H}} \mathbf{S}_{\mathrm{p}} = \mathbf{I}_{T_{\mathrm{p}}}$. For the sake of notational simplicity, in the algorithm, the training sequence matrix notation $\mathbf{S}_{\mathrm{t}}$ is used instead, where the subscript t corresponds to the number of iterations.

- Step 2 — Gradient evaluation and projection on the Riemannian parameter space: In this step, the gradient of the sum MSE cost function defined in (3.13) is determined, and the

projection on the Riemannian manifold is computed to find the descent direction and set $\mathbf{G}_t = \mathbf{D}_t$

$$\mathbf{D}_t = \mathbf{\Gamma}_t \mathbf{S}_t^{\mathrm{H}} - \mathbf{S}_t \mathbf{\Gamma}_t^{\mathrm{H}}, \tag{3.17}$$

where $\mathbf{D}_t$ is the descent direction for iteration t and Γ_t corresponds to the Euclidean gradient with respect to the training sequence matrix as given in (3.16).

- Step 3 — Stopping criteria: Check the gradient on the Riemannian parameter space (i.e., $\mathbf{D}_t$) whether it reaches convergence or not according to

$$\langle \mathbf{D}_t, \mathbf{D}_t \rangle = \frac{1}{2}\mathrm{Re}\{\mathrm{tr}(\mathbf{D}_t^{\mathrm{H}}\mathbf{D}_t)\}, \tag{3.18}$$

where Re stands for the real value. If the steepest direction is sufficiently small, which implies a close to minimum value of the cost function under consideration, the convergence is reached and the algorithm stops. The stopping criteria defined in (3.18) can be checked according to a given error tolerance ϵ, as will be given later in Section 3.5.

- Step 4 — Matrix exponential computation: Determine the local parametrisation based on the matrix exponential on the Riemannian manifold, i.e., $\mathbf{P}_t = \exp(-\mu \mathbf{D}_t)$, where μ corresponds to the step size that controls the steepest direction. The complex matrix exponential is determined by the convergent power series [119, 120], where a Matlab function *expm* is used for this purpose. An initial value for the step size is given at the beginning of the algorithm as $\mu = 1$, as considered in [119, 120]. Then, a proper step size value will be chosen according to the updating rule in step 5 and 6 called Armijo's rule [119, 120, 130]. The local parametrisation of the matrix exponential together with Armijo's step size adaptation rule would help to increase the algorithm convergence speed considerably.

- Step 5 and 6 — Step size evaluation: In every iteration t, the step size μ is required to be tuned in order to ensure an appropriate steepest direction movement towards the training solution. Two different inequalities are checked by evaluating the sum MSE cost function in order to select a proper step size value. Selecting an appropriate step size is essential because it controls the speed of convergence of the algorithm. For example,

high convergence speed may require large μ, while low steady state error would require a small μ. As such, an adaptation or self-tuning rule of μ is preferred [119]. In steps 5 and 6, the sum MSE cost function of different training sequences are evaluated based on different possibilities of μ. If μ is too small and the condition of the cost function is not met, then μ is doubled, as given in step 5. If μ is too large and the condition of the cost function is not met, then μ is halved, as provided in step 6. This step size adaptation process is very suitable for the proposed CGD algorithm since it does not require a computationally demanding calculation [119].

- Step — 7 Updating the training sequence and the gradient direction: In this step, the pilot matrix and the gradient direction are updated according to the obtained matrix exponential $\mathbf{P}_t$. Due to the unitary invariant rotation property discussed earlier in subsection 3.3.1, the obtained training sequence has unitary columns. The new conjugate gradient direction is obtained by $\mathbf{G}_{t+1} = \mathbf{D}_{t+1} + \upsilon_t \mathbf{G}_t$, where the $\mathbf{D}_{t+1}$ is determined by substituting the obtained sequence matrix into (3.17), and the parameter υ_t is given as [120]

$$\upsilon_t = \frac{\langle \mathbf{D}_{t+1} - \mathbf{D}_t, \mathbf{D}_{t+1} \rangle}{\langle \mathbf{D}_t, \mathbf{D}_t \rangle}. \tag{3.19}$$

where υ_t denotes the Polak–Ribièrre formula, as defined in [120]. Further details on expression (3.19) can be found in [120]. If the stopping criteria is met, the algorithm is executed and the training sequence is optimised. In contrast, if the stopping criteria is not met, then, a new iteration $(t + 1)$ is started. The computational complexity analysis and the simulation results for BF and RZF precoding are presented in the following section.

3.5 Computational complexity analysis and performance evaluation of different gradient-based training designs in FDD

In the first part of this section, the complexity analysis of the proposed CGD algorithm and the existing SCMI/SMSE gradient-based algorithms for DL sequence optimisation is introduced. To this end, results that characterise the overall computational complexity of the proposed CGD

Table 3.1 Computational complexity analysis

Algorithms	Complexity in flops per iteration
SCMI [53]	$T_{\mathrm{p}}^3 K + T_{\mathrm{p}}^2\big(N\,(4K-2)-1\big)\big) + 4N^2 T_{\mathrm{p}}(K-1) + NT_{\mathrm{p}}\,(7-3K) - 1$
SMSE [56]	$N\,T_{\mathrm{p}}\big(2K\,(8N-5) + 23N\,(t_d + t_h) + 80N + 13\big) - 1 + (K-1)\,X$
Proposed CGD	$NT_{\mathrm{p}} + 1/3N^2\big(T_{\mathrm{p}}(3t_{dg} + 4t_{hg} + 44) - 3\big) - 1 + (K-1)\,B$

training design and the existing SCMI/SMSE sequence designs are presented. In the second part of this section, comparisons between the achievable sum rates of the CGD training sequence design and the sequences designed based on iterative algorithms are provided. Monte Carlo simulations are considered in the experiments to obtain the results.

In the following, the *overall computational complexity* of different gradient-based iterative algorithms for DL sequence optimisation is investigated. The *overall computational complexity* is determined by multiplying the number of iterations each algorithm needs to converge by the number of flops involved per iteration, which are given in Table 3.1. For a fair comparison, the same error tolerance $\epsilon = 0.001$ for all the training sequence designs is used. Table 3.1 summarises the complexity analysis based on the number of multiplications and additions in flops [123] per iteration for the three training sequence designs considered. Parameters (t_{dg}, t_{hg}) and (t_d, t_h) represent the number of iterations required for doubling and halving the step size in the proposed CGD method and the SMSE iterative algorithm, respectively. The subscript dg denotes doubling the step size in the proposed CGD method, whereas subscript hg refers to the halving the step size in the proposed CGD, where the letter g stands for the gradient. The variable X is given as $X = \big(5T_{\mathrm{p}}^3 + 2T_{\mathrm{p}}^2(7N-1)\big)$ for SMSE and the variable B is given as $B = \big(5T_{\mathrm{p}}^3 + 16N^2T_{\mathrm{p}} + 14NT_{\mathrm{p}}^2 - 2T_{\mathrm{p}}^2 - 10NT_{\mathrm{p}}\big)$ for CGD. Below are the essential details of the flops calculation (obtained by counting the number of multiplication and additions per iteration) for the proposed CSD algorithm. Appendix A explains further how the computational complexity of different iterative algorithms were obtained in Table 3.1.

- The complexity of calculating the gradient in step 2 of Algorithm 1, so that the term $\sum_{k=1}^{K} \mathbf{R}_k^2\mathbf{S}_{\mathrm{p}}\big(\mathbf{S}_{\mathrm{p}}^{\mathrm{H}}\mathbf{R}_k\mathbf{S}_{\mathrm{p}} + \mathbf{I}_{T_{\mathrm{p}}}\big)^{-1} + \mathbf{R}_k\mathbf{S}_{\mathrm{p}} \times \big(\mathbf{S}_{\mathrm{p}}^{\mathrm{H}}\mathbf{R}_k\mathbf{S}_{\mathrm{p}} + \mathbf{I}_{T_{\mathrm{p}}}\big)^{-1}\mathbf{S}_{\mathrm{p}}^{\mathrm{H}}\,\mathbf{R}_k^2\mathbf{S}_{\mathrm{p}}\big(\mathbf{S}_{\mathrm{p}}^{\mathrm{H}}\mathbf{R}_k\mathbf{S}_{\mathrm{p}} + \mathbf{I}_{T_{\mathrm{p}}}\big)^{-1}$ requires $(K-1)\big(5T_{\mathrm{p}}^3 + 16N^2T_{\mathrm{p}} + 14NT_{\mathrm{p}}^2 - 2T_{\mathrm{p}}^2 - 10NT_{\mathrm{p}}\big)$ flops.

- Calculating the matrix exponential requires $2(2/3)N^2T_\mathrm{p}$ flops [119, 120].
- Checking the gradient convergence in step 3 using the squared Frobenius norm of an $N \times T_\mathrm{p}$ matrix requires $2NT_\mathrm{p} - 1$ flops [123].
- Calculating the term step size adaptation in step 5 and step 6 of Algorithm 1 requires $\big(2(2/3)N^2T_\mathrm{p} + (t_d + 3)N^2T_\mathrm{p} + t_h(2(2/3)N^2T_\mathrm{p}) + 3N^2T_\mathrm{p}\big)$ flops [119].
- Multiplying an $N \times N$ matrix with an $N \times T_\mathrm{p}$ matrix, entails $2N^2T_\mathrm{p} - NT_\mathrm{p}$ flops.
- Multiplying a $T_\mathrm{p} \times N$ matrix with an $N \times T_\mathrm{p}$ matrix, entails $2NT_\mathrm{p}^2 - T_\mathrm{p}^2$ flops.
- Multiplying a $T_\mathrm{p} \times N$ matrix with an $N \times N$ matrix, entails $2N^2T_\mathrm{p} - NT_\mathrm{p}$ flops.
- Multiplying a $T_\mathrm{p} \times T_\mathrm{p}$ matrix with a $T_\mathrm{p} \times T_\mathrm{p}$ matrix requires $2T_\mathrm{p}^3 - T_\mathrm{p}^2$ flops.
- Multiplying an $N \times N$ matrix with an $N \times N$ matrix, entails $2N^3 - N^2$ flops.
- Multiplying an $N \times T_\mathrm{p}$ matrix with a $T_\mathrm{p} \times T_\mathrm{p}$ matrix, entails $2NT_\mathrm{p}^2 - NT_\mathrm{p}$ flops.
- Multiplying an $N \times T_\mathrm{p}$ matrix with a $T_\mathrm{p} \times N$ matrix, entails $2N^2T_\mathrm{p} - N^2$ flops.
- The scalar matrix multiplication with an $N \times T_\mathrm{p}$ matrix needs NT_p flops.
- Adding an $N \times T_\mathrm{p}$ matrix to an $N \times T_\mathrm{p}$ matrix requires NT_p flops.
- Inverting of a $T_\mathrm{p} \times T_\mathrm{p}$ matrix requires T^3 flops.

3.5.1 Achievable sum rate performance evaluation of different gradient-based algorithms in FDD massive MIMO

In this subsection, the achievable sum rate performance of the proposed CGD training design and the state-of-the-art SCMI/SMSE sequence designs are compared. The curves for the sum rate with the BF and RZF precoders are plotted based on equations (2.3) and (3.5). The number of incidences of the channel realisations used in the Monte carlo simulations is 10^4. The results in Fig. 3.1 demonstrate that the proposed CGD training design achieves the same rate performances as state-of-the-art training sequence designs. The results in Fig. 3.1 confirm

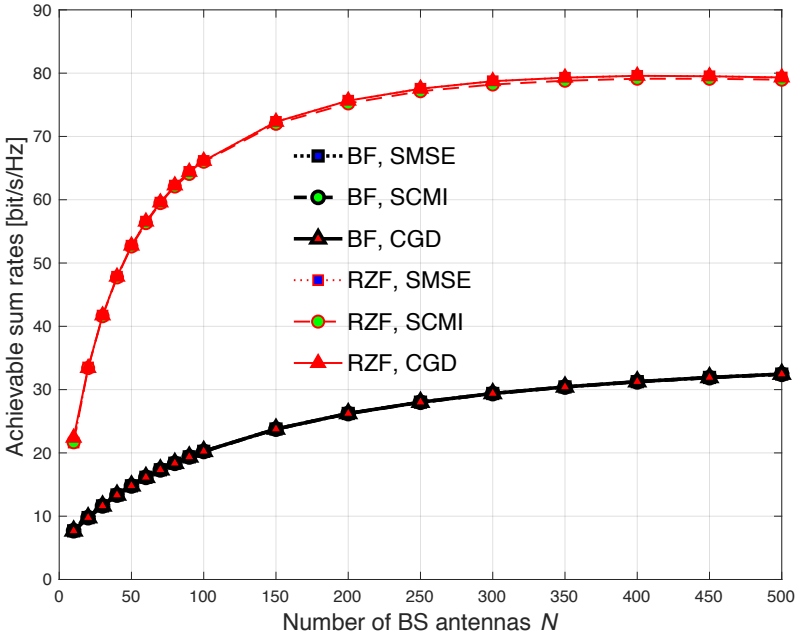


Figure 3.1 Achievable sum rate versus number of BS antennas N, comparing different training sequence designs using the OR model with $\omega_{\mathrm{H}} = 2.5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 10$ users.

that significant improvement in the sum rate performances are obtained for both the BF and RZF precoders when the channels are strongly correlated.

Fig. 3.2 examines the achievable sum rate comparing different training sequence designs under a weak channel correlation, i.e., $\omega_{\mathrm{H}} = 25°$, $D = 1$. The other salient system parameters remain unchanged at $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 10$ users. Fig. 3.2 shows that, with weak channel correlation, some loss in the rate performances are obtained for both the BF and RZF precoders with the SCMI training sequence design, which uses a Lagrangian multiplier iterative algorithm, in comparison with the SMSE and the proposed CGD training designs. The degradation in the sum rate performance with the SCMI training design can be justified as the Lagrangian-based iterative algorithm dose not consider the spatial properties of the parameter space such as a Grassmannian space or Riemannian manifold in the training sequences optimisation. In particular, using the smooth parameter space in the training sequence optimisation allows an efficient precoding design to be achieved, and hence maximising the sum rate performance. Furthermore, Fig. 3.2 demonstrates that feasible sum rate performances are obtained in all the training methods when the channels are, relatively, weak correlated (i.e., $\omega_{\mathrm{H}} = 25°$, $D = 1$). This is due to that fact that the achievable sum rate is enhanced by optimising the

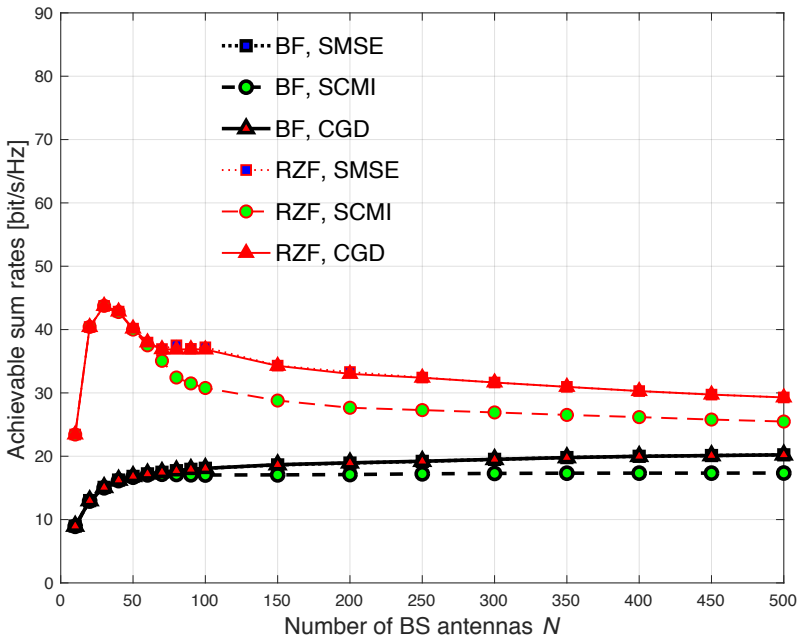


Figure 3.2 Achievable sum rate versus number of BS antennas N, comparing different training sequence designs using the OR model with $\omega_{\mathrm{H}} = 25°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 10$ users.

pilot length based on the maximisation of $(1 - T_{\mathrm{p}}/T_{\mathrm{c}}) \sum_{k=1}^{K} \log_2(1 + \mathrm{SINR}_k)$, where the pre-log fraction is accounted for.

Fig. 3.3 investigates the sum rate performance comparing different training sequence designs when the number of users is increased, i.e., $K = 12$. The results in Fig. 3.3 show that the peak sum rate performances obtained for both the BF and RZF precoders increase when K increases, as expected. While the results provided in this chapter are based on Monte Carlo simulations, random matrix theory methods will be used in Chapter 5 to accurately approximate the achievable sum rate performance of the BF and RZF precoders in an FDD massive MIMO system. Overall, the results provided in this subsection show that almost the same rate performances are obtained in all the training sequence designs considered when the channels are strongly correlated.

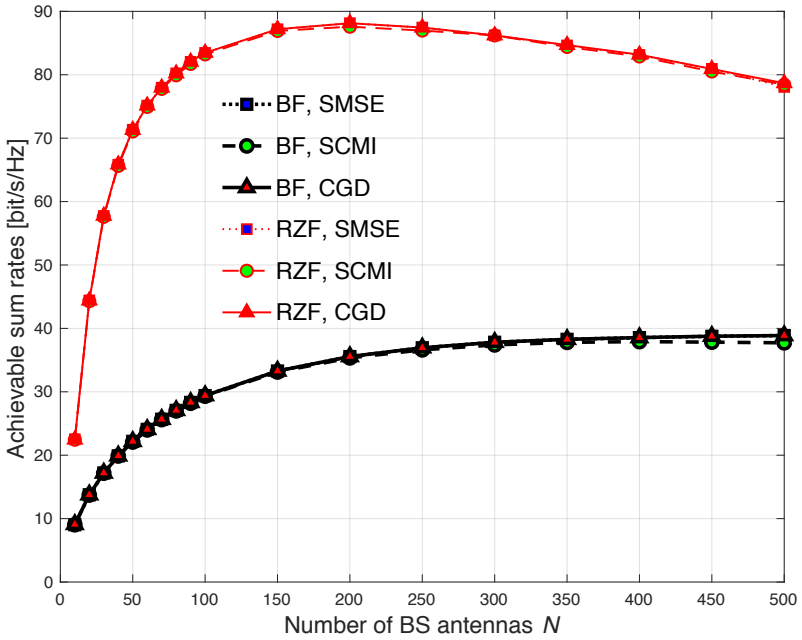
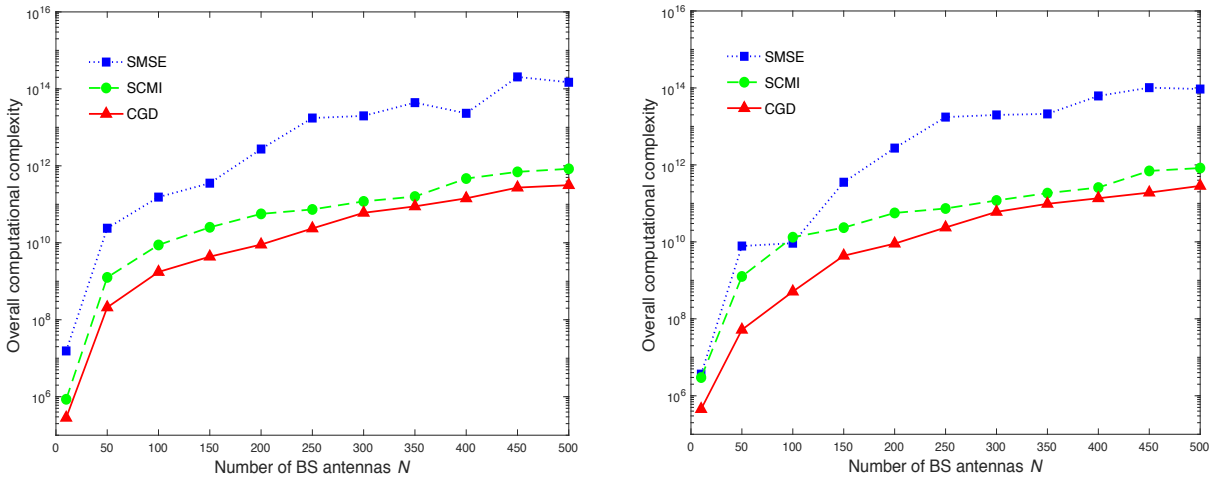


Figure 3.3 Achievable sum rate versus number of BS antennas N, comparing different training sequence designs using the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 12$ users.



(a) BF precoding.

(b) RZF precoding.

Figure 3.4 Overall computational complexity versus the number of BS antennas N comparing different training sequence methods corresponding to the pilot length in BF and RZF precoding in the OR model with $\omega_{\mathrm{H}} = 2.5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 10$ users.

Table 3.2 The number of iterations per each training sequence algorithm requires to converge with respect to the number of BS antennas N for the pilot length in BF and RZF precoding in the OR model with $\omega_{\mathrm{H}} = 2.5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 10$ users.

N	SCMI	SMSE	CGD
10	85	120	66
50	3151	767	201
100	4593	786	226
150	4210	669	190
200	4120	1276	182
250	2816	2172	223
300	2915	1909	315
350	2655	2341	316
400	5557	1419	343
450	5727	3766	420
500	5245	2637	379

(a) BF precoding.

N	SCMI	SMSE	CGD
10	151	40	66
50	3151	413	54
100	5660	166	84
150	3376	669	190
200	4120	1276	182
250	2816	2172	223
300	2915	1909	315
350	2849	1562	316
400	2851	2261	309
450	5727	2500	326
500	4893	2022	336

(b) RZF precoding.

3.5.2 Computational complexity evaluation of different gradient-based training designs for FDD

The results from Fig. 3.4 to Fig. 3.8 show plots of the overall computational complexity versus the number of BS antennas N, comparing the proposed CGD training sequence design with SMSE and SCMI training designs using the pilot length for BF and RZF precoding. The number of iterations each algorithm required to converge are provided in Table 3.2 to Table 3.6, which are used to obtain the simulated results from Fig. 3.4 to Fig. 3.8, respectively. The results are presented based on the scattering OR channel model, described in Chapter 2 subsection 2.7.2, when the system parameters K, N, T_{c}, SNR, angle of arrivals (AoAs), normalised antenna spacing D, and angular spread ω_{H} are used. Unless otherwise specified, the AoAs are distributed in the range $[-60°, 60°]$, i.e., $\theta_k = \{-57.5°, -45°, -41.5°, -23°, -7.5°, 7.5°, 23.5°, 41.5°, 45°, 57.5°\}$, [53, 131].

The results in Fig. 3.4 and Fig. 3.5 were obtained with $\omega_{\mathrm{H}} = 2.5°$, $D = 1/2$ and $\omega_{\mathrm{H}} = 25°$, $D = 1$, respectively, while the other system parameters are $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 10$ users. Note that parameters $\omega_{\mathrm{H}} = 2.5°$, $D = 1/2$ corresponds to strong spatial correlation whereas $\omega_{\mathrm{H}} = 25°$, $D = 1$ imply relatively weak correlation. The results in Fig. 3.6 were obtained with SNR = 5 dB, while the other system parameters are $T_{\mathrm{c}} = 100$

symbols, $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, and $K = 10$ users. The results in Fig. 3.7 and Fig. 3.8 were obtained with different number of users; $K = 6$ and $K = 12$, respectively, while the other system parameters are $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, $\omega_{\mathrm{H}} = 5°$, and $D = 1/2$. The parameter values of AoAs are chosen with $K = 6$, $\theta_k = \{-41.5°, -23°, -7.5°, 7.5°, 23.5°, 41.5°\}$, while for $K = 12$, $\theta_k = \{-57.5°, -45°, -41.5°, -30°, -23°, -7.5°, 7.5°, 23.5°, 30°, 41.5°, 45°, 57.5°\}$. These salient system parameters are frequently encountered in the open literature on massive MIMO evaluation.

The results show that the proposed CGD iterative method for the DL training sequence optimisation achieves a faster convergence rate (as illustrated by the number of iterations given in Table 3.2 to Table 3.6 including the number of iterations required for doubling and halving the step size), compared with the state-of-the-art iterative algorithms. The fast convergence speed provided by the proposed CGD algorithm arises from the use of the matrix exponential search on the Riemannian manifold. Specifically, the proposed CGD iterative algorithm exhibits fewer number of iterations to converge due to the significant advantages of the almost periodic property of the matrix exponential search on the smooth Riemannian manifold. Although the proposed CGD provides fast convergence speed, the results demonstrate that as the number of BS antennas N increases, the SCMI and proposed CGD iterative algorithms exhibit almost the same overall computational complexity. For example, the results in Fig. 3.4 show that at $N = 350$, the proposed CGD iterative algorithm offers comparable overall computational complexity to the SCMI algorithm, which implies that the SCMI iterative algorithm requires fewer flops at this BS array size. Nonetheless, the proposed CGD iterative algorithm is still provided the a larger sum rate performance than the SCMI algorithm under a relatively weak correlated channels, see e.g. Fig. 3.2. Overall, the results provided in this subsection indicate that the relative ratio between the three iterative algorithms for the training sequence optimisation depends on the number of BS antennas N and the levels of spatial correlation.

Table 3.3 The number of iterations per each training sequence algorithm requires to converge with respect to the number of BS antennas N for the pilot length in BF and RZF precoding in the OR model with $\omega_{\mathrm{H}} = 25°$, $D = 1$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 10$ users.

N	SCMI	SMSE	CGD
10	316	25	23
50	13742	394	129
100	15806	1278	162
150	8449	1606	217
200	10297	1392	311
250	7069	2298	310
300	6401	2437	332
350	7360	1958	313
400	7897	2294	317
450	6484	2264	278
500	9592	2106	358

(a) BF precoding.

N	SCMI	SMSE	CGD
10	163	3	3
50	5865	136	64
100	11823	870	270
150	10546	1416	238
200	7824	844	309
250	7069	2127	267
300	7568	2833	271
350	6618	2563	287
400	6959	2655	273
450	6448	2483	296
500	9920	2483	292

(b) RZF precoding.

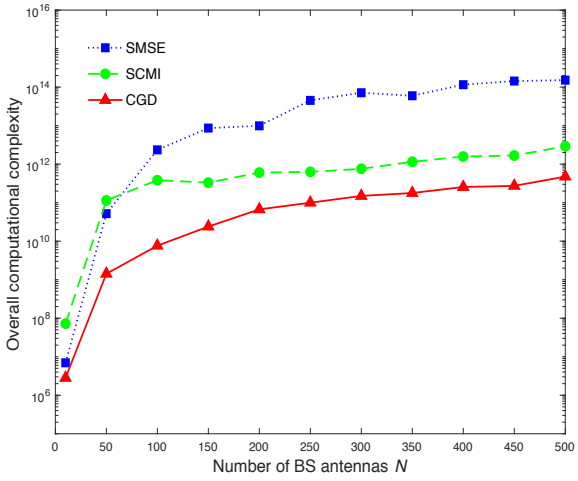
Table 3.4 The number of iterations per each training sequence algorithm requires to converge with respect to the number of BS antennas N for the pilot length in BF and RZF precoding in the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 5 dB and $K = 10$ users.

N	SCMI	SMSE	CGD
10	62	119	100
50	2589	866	258
100	3189	602	165
150	2417	836	272
200	3485	822	354
250	4161	3028	269
300	4787	3078	342
350	4455	3110	473
400	5933	2384	544
450	6808	2800	583
500	6156	1803	508

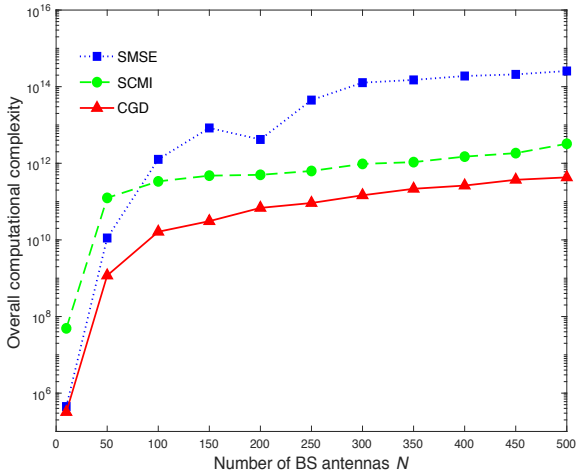
(a) BF precoding.

N	SCMI	SMSE	CGD
10	253	56	71
50	2465	337	106
100	2044	602	165
150	1979	836	272
200	3485	1583	268
250	3711	3028	269
300	4787	1298	338
350	4833	884	358
400	4845	2384	470
450	6451	1338	494
500	8930	3277	553

(b) RZF precoding.

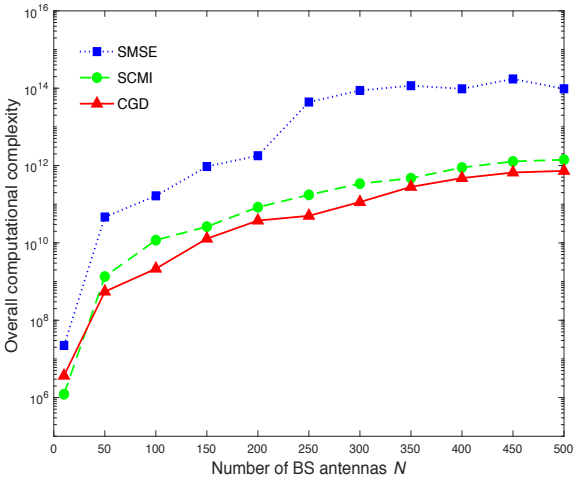


(a) BF precoding.

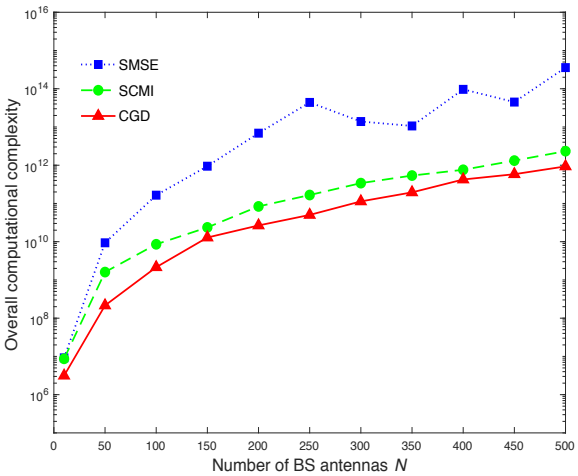


(b) RZF precoding.

Figure 3.5 Overall computational complexity versus the number of BS antennas N comparing different training sequence methods corresponding to the pilot length in BF and RZF precoding in the OR model with $\omega_{\mathrm{H}} = 25°$, $D = 1$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 10$ users.



(a) BF precoding.



(b) RZF precoding.

Figure 3.6 Overall computational complexity versus the number of BS antennas N comparing different training sequence methods corresponding to the pilot length in BF and RZF precoding in the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 5 dB and $K = 10$ users.

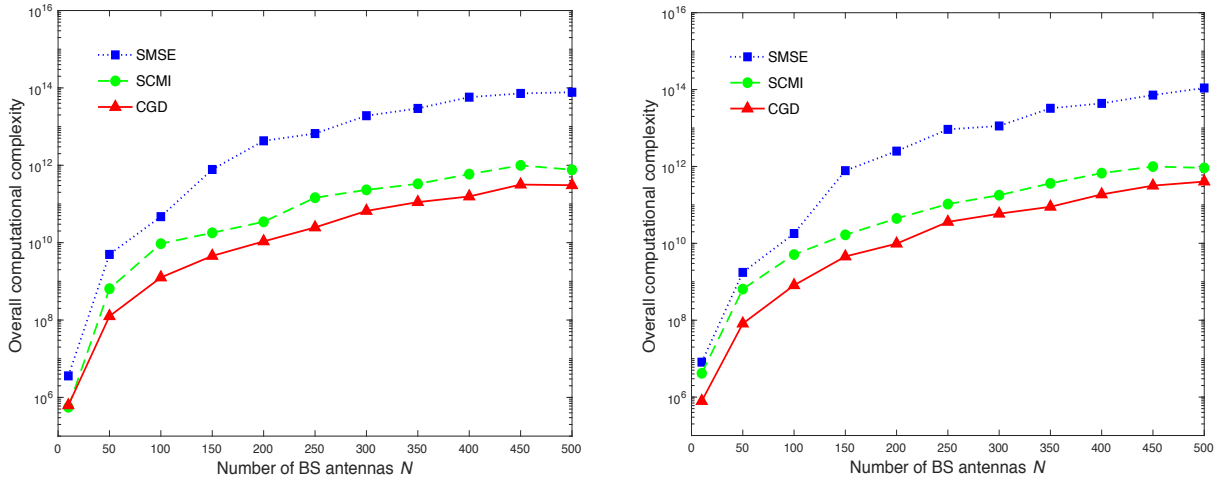
Table 3.5 The number of iterations per each training sequence algorithm requires to converge with respect to the number of BS antennas N for the pilot length in BF and RZF precoding in the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 6$ users.

N	SCMI	SMSE	CGD
10	49	55	50
50	1763	288	117
100	4545	318	140
150	2689	747	155
200	2391	1091	165
250	5800	1041	200
300	5831	1355	280
350	5627	1428	298
400	7380	1763	303
450	9942	1739	403
500	6266	1633	347

(a) BF precoding.

N	SCMI	SMSE	CGD
10	209	60	24
50	1763	155	73
100	2177	188	102
150	2281	747	155
200	2679	853	161
250	4789	1225	264
300	4776	1126	263
350	6494	1580	263
400	8822	1490	346
450	9942	1739	403
500	7171	1917	400

(b) RZF precoding.



(a) BF precoding.

(b) RZF precoding.

Figure 3.7 Overall computational complexity versus the number of BS antennas N comparing different training sequence methods corresponding to the pilot length in BF and RZF precoding in the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 6$ users.

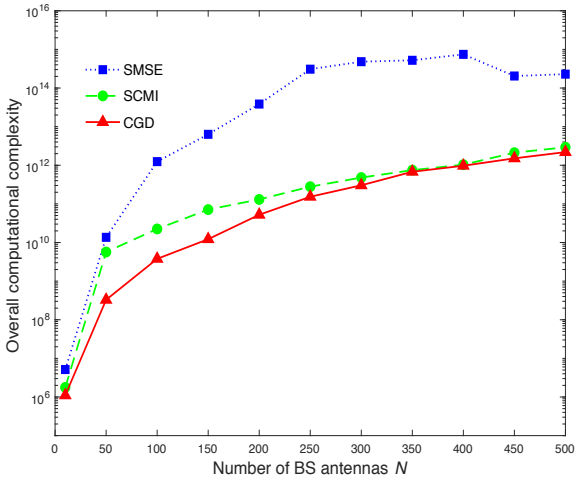
Table 3.6 The number of iterations per each training sequence algorithm requires to converge with respect to the number of BS antennas N for the pilot length in BF and RZF precoding in the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 12$ users.

N	SCMI	SMSE	CGD
10	74	49	47
50	7154	469	163
100	4979	1574	234
150	4882	2149	237
200	4139	3654	385
250	5099	7357	533
300	5539	7101	616
350	5508	6428	785
400	5409	6772	827
450	8468	3089	886
500	9564	2901	974

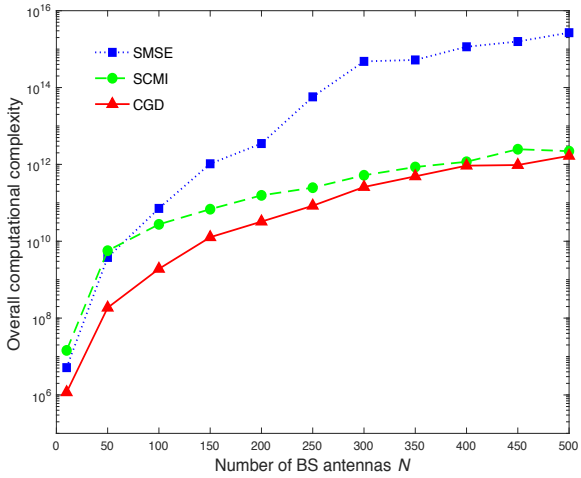
(a) BF precoding.

N	SCMI	SMSE	CGD
10	348	49	26
50	7154	220	88
100	5385	348	135
150	4291	836	238
200	4321	1001	250
250	3796	3087	329
300	5686	7113	537
350	6008	6428	584
400	5850	7133	761
450	10823	6915	588
500	7536	7736	703

(b) RZF precoding.



(a) BF precoding.



(b) RZF precoding.

Figure 3.8 Overall computational complexity versus the number of BS antennas N comparing different training sequence methods corresponding to the pilot length in BF and RZF precoding in the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 12$ users.

3.6 Summary

Chapter 2 explained that optimising the DL training sequences for CSI estimation in a general scenario with single-stage precoding and K independent channel correlation matrices is very

challenging. Although advanced gradient-based iterative algorithms have been developed to address this challenge, they exhibit slow convergence speed. To overcome this issue, this chapter investigates a geometrical optimisation approach, which utilises a computationally efficient Riemannian manifold, with an aim to maximise the sum rate performance, while improving the convergence rate to provide a fast DL CSI estimation for the FDD massive MIMO systems. To this end, a new CGD algorithm is proposed, which uses the matrix exponential search on the Riemannian manifold.

The sum rate performance of the proposed CGD algorithm is also compared with the rates achieved by the state-of-the-art iterative algorithms. In order to examine the performance of the proposed CGD algorithm for DL training sequence optimisation, comprehensive simulated results are presented in this chapter. Additionally, the computational complexity analysis and the number of convergence iterations of the proposed CGD algorithm and the state-of-the-art SCMI/SMSE iterative algorithms are provided. Consequently, the overall computational complexity of the proposed CGD training design is compared with the state-of-the-art iterative algorithms. Importantly, the results demonstrated that, under a relatively weak correlate channels, over a 5bit/s/Hz improvement in the achievable sum rate is obtained using the proposed CGD iterative algorithm in comparison to the SCMI algorithm. This achievement signifies the advantages of the proposed approach for the training sequence optimisation of FDD massive MIMO systems compared with the SCMI approach. This comparison is carried out based on the scattering OR channel model. The results indicated that the proposed CGD algorithm is able to achieve a best sum rate performance owing to a faster convergence rate. This is attributed to the efficient matrix exponential search on the Riemannian manifold, which allows the algorithm to converge faster with a fewer number of iterations. The results showed that all the training sequence designs considered exhibit essentially the same sum rate performances in both BF and RZF precoding.

While the iterative algorithms introduced in this chapter are highly efficient in terms of optimising the DL training sequences to achieve the best sum rate performances for FDD massive MIMO systems, the results demonstrated that they require hundreds or thousands of iterations to reach convergence, which is still overly complex. This could be a bottleneck in real-time system where the CSI needs to be estimated more frequently to track the user channels variation,

especially when a scenario with limited coherence time is considered. In particular, by the time the training sequence solution is obtained based on such iterative approaches, the user channels may have changed, and thus, the obtained CSI solution would be outdated. This prompts the motivation for research into developing a non-iterative approach for DL training sequence design and CSI estimation in an FDD massive MIMO system in order to satisfy both complexity and latency practical implementation concerns. Furthermore, at this point in the thesis, the optimum training sequence length aiming to maximise the achievable sum rate performance is obtained by exhaustively searching from all possible combinations of $1 \leq T_{\mathrm{p}} \leq \min\left(N, T_{\mathrm{c}}\right)$. However, with the scaling up of the BS array and increasing of the coherence time, such an exhaustive search might be time and resource consuming, and hence, practically infeasible.

The next chapter will investigate a non-iterative approach for the DL training sequence design in an FDD massive MIMO system. Additionally, the focus will be on developing a tractable framework analysis to find a closed-form solution for the optimum pilot length in order to obviate the need for executing an exhaustive search approach.

Chapter 4

Training Sequence Design in FDD Massive MIMO Based on a Common Spatial Correlation Between Users

4.1 Introduction

Chapter 3 showed how the DL training sequences are optimised using gradient-based iterative algorithms. Such iterative approaches provided useful characterisation of the best sum rate performances that can be achieved in an FDD massive MIMO system with single-stage precoding and distinct spatial channel correlations. However, the practical implementation of these iterative algorithms is limited by their complexity: many iterations are required to reach convergence and each iteration needs intensive computing resources, due to the inversion and multiplication of large matrices. As such, the convergence time of these iterative approaches are overwhelming. In addition, in practical scenario with limited coherence time, the CSI estimation must be determined more often, and thus, iterative-based solutions for the DL training sequence design may be infeasible. Therefore, identifying non-iterative algorithms for the DL training sequences that are less computationally demanding, through reduced design complexity, is a primary objective of this thesis. Furthermore, finding the optimum training sequence length analytically in closed-form with limited coherence time is of interest. Note that minimising the training

duration maximises the achievable throughput, especially in channels with limited coherence time, as explained in Chapter 2.

Develop a low complexity and tractable training solution. The computational complexity of the proposed training design is quantified and compared with Chapter 3. Chapter 4 is essential since it establishes the analytical framework for the optimum training sequence and for superposition. Chapter 4 is special case of superposition in Chapter 5 since the superposition is reduced to one training case for all users. Chapter 4 derives a novel closed-form solution of the optimum pilot length. Still a special case, which doesn't cover the more practical scenario where covariance matrices are different, hence Chapter 5 is needed.

4.1.1 Contributions and chapter findings

This chapter develops a low complexity training solution for DL CSI estimation in an FDD massive MIMO communication system with single-stage precoding in a scenario where all users exhibit the same spatial channel correlation $\mathbf{R}_k = \mathbf{R}\ \forall k = 1,\ldots,K$ at the BS. In addition, this chapter establishes a tractable framework analysis for the optimum training sequence that maximises the achievable sum rate in FDD massive MIMO systems with limited coherence time. Specifically, a non-iterative training design is developed based on sequences constructed from the effective directions (eigenvectors) of the transmit correlation matrix, which are the same for all users. Further, this chapter derives a novel analytical closed-form solution for the optimum pilot length that maximises the achievable sum rate in FDD massive MIMO systems. This analytical solution would obviate the need for exhaustive searching to find the optimum pilot length that was necessary in Chapter 3. Note that an exhaustive search is time and resource consuming that might be practically unsuitable, especially for $N, K, T_{\mathrm{c}} \gg 1$. In Chapter 5, the tractable framework analysis for both the optimum pilot structure and the optimum training duration, which are established in this chapter, will be built upon and extended to a more general scenario in which the users exhibit different spatial correlations. An analytical method based on a large-dimension random matrix theory, along with the P-DoF channel model is developed in this chapter in order to provide a straightforward procedure for the achievable sum rate optimisation. This enables the achievable sum rate analysis of the BF and RZF precoders in an FDD massive MIMO system under consideration to be expressed in a closed-form, so that the

numerical evaluation can be readily calculated without resorting to computationally demanding Monte Carlo simulations. Results characterising the system performance of the BF and RZF precoders are provided for the P-DoF channel model and the scattering OR channel model. To gain further insight into the impact of training sequence length on the sum rate performance, comparison between the sum rate of perfect and imperfect channel estimation is provided. The computational complexity of the training sequence design for an massive MIMO system under consideration is quantified and compared with the state-of-the-art iterative approaches. The results indicate that a feasible sum rate performance can be achieved when finite length pilot sequences are used in the DL CSI estimation of an FDD massive MIMO system. The finding in this chapter is supported by a rigorous mathematical analysis that agrees with the simulated results, as desired.

4.2 Achievable sum rate analysis based on perfect channel estimation

To provide a baseline for comparison with an imperfect channel estimation for the massive MIMO system under consideration, this section explores the achievable sum rate analysis of the BF and RZF precoders with perfect knowledge of the channel. A large system limit based on random matrix theory methods, as explained in Chapter 2 subsection 2.5.2, is adopted to provide an analytical asymptotic approximation of the SINR expression in (3.5), and hence the sum rate, for the BF and RZF precoders. A DL single-cell massive MIMO communications systems, which is introduced in Chapter 3 section 3.2, is considered. Accordingly, the DL achievable sum rate C^{PF}, where PF indicates a prefect channel estimation, based on a common correlation and when the BS employs precoders with perfect channel estimation can be expressed as

$$\mathrm{C}^{\mathrm{PF}} = K \cdot \log_2\big(1 + \mathrm{SINR}\big), \tag{4.1}$$

where SINR is the effective SINR defined in (3.5). Note that the estimates $\hat{\mathbf{H}}$ of the DL channel, defined in (3.6) is now replaced with the DL channel $\mathbf{H}$. As such, the BF and RZF precoders

are now designed based on perfect channel estimation. Comparing with the SINR term in (3.5) presented in Chapter 3, the SINR expression in (4.1) is now missing the subscript k. This is due to the condition $\mathbf{R}_k = \mathbf{R}\ \forall k = 1, \ldots, K$, which leads to statistically the same SINR term for all users. The associated DL achievable sum rate with perfect channel estimation, denoted by $\bar{\mathrm{C}}^{\mathrm{PF}}$, where the notation PF refers to a prefect channel estimation, for the massive MIMO system with common spatial correlation between users is given by

$$\bar{\mathrm{C}}^{\mathrm{PF}} = K \cdot \log_2\big(1 + \overline{\mathrm{SINR}}\big), \tag{4.2}$$

where $\overline{\mathrm{SINR}}$ is an asymptotically tight approximation for the average effective SINR in (3.5). Consistent with previous research work in [30, 32, 132], the numerical results provided in this chapter indicate that these asymptotic approximations are also highly valid for practical, finite values of N and K.

The following Propositions (i.e., Propositions 4.1 and 4.2) provide an asymptotic analysis results related to the SINR of the BF and RZF precoders with general correlation model, based on random matrix theory methods [30]. In particular, Propositions 4.1 and 4.2 are modified versions of Theorem 4 and Theorem 6 in [30], respectively, where the system under consideration is carried out on the downlink of a single-cell FDD massive MIMO system and when the BS employs precoding with perfect channel estimation. Additionally, Propositions 4.1 and 4.1 consider a special scenario where users exhibit the same spatial correlation at the BS.

Proposition 4.1. *let* $\mathrm{SINR}^{\mathrm{BF,PF}}$ *denote the* SINR *for BF precoding as defined in* (3.5) *with perfect channel estimation, where the subscript* PF *refers to the perfect channel estimation. An asymptotically tight approximation for the* $\mathrm{SINR}^{\mathrm{BF,PF}}$ *for the regime where N and K go to infinity with a given fixed ratio reads*

$$\overline{\mathrm{SINR}}^{\mathrm{BF,PF}} = \frac{\mathrm{tr}(\mathbf{R})}{\frac{1}{\rho_{\mathrm{d}}} + K\big(\mathrm{tr}(\mathbf{R})\big)^{-1}\mathrm{tr}(\mathbf{R}^2)}, \tag{4.3}$$

where ρ_{d} *is the average per-user BS transmit power during the data-phase and* $\mathbf{R}$ *is the transmit channel covariance matrix at the BS-side, which is considered to be the same for all the users.*

The SINR expression for the BF precoder in (4.3) is generally valid for any channel correla-

tion model. A further simplified expression related to the SINR term, and hence the achievable sum rate performance, for the BF precoder is provided later in subsection 4.2.1 based on the P-DoF channel model. Unlike the SINR expression of BF precoding, the SINR expression of the RZF precoder is provided in terms of several auxiliary variables. These auxiliary variables arise from the asymptotic random matrix theory methods [30] and, in general, require to be numerically solved before the SINR approximation is calculated. Nonetheless, a simplified analysis for the RZF precoder with perfect channel estimation and common correlation between users is developed in subsection 4.2.1 based on the P-DoF channel model. The following Proposition elaborates the main result related to the SINR analysis for the RZF precoder under a general channel correlation model and when the users exhibit the same spatial correlation at the BS-side.

Proposition 4.2. *Let* $\mathrm{SINR}^{\mathrm{RZF,PF}}$ *denote the* SINR *of RZF precoding as defined in* (3.5) *with perfect channel estimation, where* PF *denotes the perfect channel estimation. An asymptotically tight approximation for the* $\mathrm{SINR}^{\mathrm{RZF,PF}}$ *for an FDD massive MIMO system with a common correlation between users is provided as*

$$\overline{\mathrm{SINR}}^{\mathrm{RZF,PF}} = \frac{N\,\ddot{\xi}\,\ddot{\delta}^2}{\frac{(1+\ddot{\delta})^2}{\rho_{\mathrm{d}}K} + K\ddot{\xi}\ddot{\mu}}, \tag{4.4}$$

where the parameter $\ddot{\xi} \in \mathbb{R}$ *is defined later in* (4.10)*. Defining a recursion on integer* t*, where* $t = 1, 2, \ldots,$

$$\ddot{\delta}^{(t)} = \frac{1}{N}\mathrm{tr}\left(\mathbf{R}\left(\zeta\mathbf{I}_N + \frac{K\mathbf{R}}{N\left(1+\ddot{\delta}^{(t-1)}\right)}\right)^{-1}\right), \tag{4.5}$$

with an initial value $\ddot{\delta}^{(0)} = 1/\zeta$*, where* $\zeta > 0$ *is the regularisation parameter of the RZF precoder, which is considered to be the inverse of the per-user SNR,* $1/\rho_{\mathrm{d}}$ *as given in [30], and the auxiliary variable* $\ddot{\delta} \in \mathbb{R}$ *is a solution to a fixed-point algorithm, where the solution to* $\ddot{\delta}$ *is obtained numerically as the limit*

$$\ddot{\delta} = \lim_{t\to\infty} \ddot{\delta}^{(t)}. \tag{4.6}$$

A unique solution to a fixed-point algorithm defined in (4.5) *can be obtained analytically in a closed-form using the P-DoF channel model, as developed in Proposition 4.2 (see subsection 4.2.1). When the convergence is reached in* (4.6)*, a solution for* $\ddot{\delta}$ *is obtained numerically,*

which is then substituted into (4.7)

$$\ddot{\mathbf{T}} = \left(\zeta \mathbf{I}_N + \frac{K\mathbf{R}}{N\left(1+\ddot{\delta}\right)}\right)^{-1}, \tag{4.7}$$

to obtain the auxiliary random matrix $\ddot{\mathbf{T}} \in \mathbb{C}^{N\times N}$*. Matrix variable* $\ddot{\ddot{\mathbf{T}}} \in \mathbb{C}^{N\times N}$ *is given by*

$$\ddot{\ddot{\mathbf{T}}} = \ddot{\mathbf{T}}\left(\mathbf{I}_N + \frac{K\mathbf{R}\ddot{\ddot{\delta}}}{N\left(1+\ddot{\delta}\right)^2}\right)\ddot{\mathbf{T}}, \tag{4.8}$$

where the auxiliary parameter $\ddot{\ddot{\delta}} \in \mathbb{R}$ *is obtained as*

$$\ddot{\ddot{\delta}} = \left(1 - \frac{\mathrm{tr}\left(\mathbf{R}\ddot{\mathbf{T}}\mathbf{R}\ddot{\mathbf{T}}\right)}{\left(N(1+\ddot{\delta})\right)^2}\right)^{-1}\frac{1}{N}\mathrm{tr}\left(\ddot{\mathbf{T}}\mathbf{R}\ddot{\mathbf{T}}\right). \tag{4.9}$$

The auxiliary variable $\ddot{\xi} \in \mathbb{R}$ *in* (4.4) *is the precoder normalisation and given by substituting* $\ddot{\mathbf{T}}$ *and* $\ddot{\ddot{\mathbf{T}}}$ *into*

$$\ddot{\xi} = \left(\mathrm{tr}\left(\ddot{\mathbf{T}}\right) - \zeta\mathrm{tr}\left(\ddot{\ddot{\mathbf{T}}}\right)\right)^{-1}. \tag{4.10}$$

The auxiliary variable $\ddot{\mu} \in \mathbb{R}$ *is determined based on the dominated convergence and the continuous mapping theorems as developed in [30], which is modified here for an FDD massive MIMO system under consideration, and hence, it is provided from the expressions in* (4.11)*,* (4.12) *and* (4.13)*.*

$$\ddot{\mu} = \frac{1}{N}\mathrm{tr}\left(\mathbf{R}\ddot{\mathbf{T}}'\right) - \frac{2\mathrm{Re}\left(\mathrm{tr}\left(\mathbf{R}\ddot{\mathbf{T}}\right)\mathrm{tr}\left(\mathbf{R}\ddot{\mathbf{T}}'\right)\right)\left(1+\ddot{\delta}\right) - \mathrm{tr}\left(\mathbf{R}\ddot{\mathbf{T}}\right)^2\ddot{\delta}'}{\left(N(1+\ddot{\delta})\right)^2} \tag{4.11}$$

$$\ddot{\mathbf{T}}' = \ddot{\mathbf{T}}\left(\mathbf{R} + \frac{K\mathbf{R}\ddot{\delta}'}{N\left(1+\ddot{\delta}\right)^2}\right)\ddot{\mathbf{T}} \tag{4.12}$$

$$\ddot{\delta}' = \left(1 - \frac{\mathrm{tr}\left(\mathbf{R}\ddot{\mathbf{T}}\mathbf{R}\ddot{\mathbf{T}}\right)}{\left(N\left(1+\ddot{\delta}\right)\right)^2}\right)^{-1}\frac{1}{N}\mathrm{tr}\left(\ddot{\mathbf{T}}\mathbf{R}\ddot{\mathbf{T}}\mathbf{R}\right) \tag{4.13}$$

The SINR expression in (4.4) for the RZF precoder is generally valid for any channel correlation model. A simplified closed-form expression of the SINR for the BF and RZF precoders

with perfect channel estimation is provided in the next subsection using the P-DoF channel model.

4.2.1 Sum rate analysis of the BF and RZF precoders with perfect channel estimation based on the P-DoF channel model and common correlation between users

In this subsection, an analytical P-DoF channel model, described in Chapter 2 section 2.7.1, is used to obtain a closed-form analysis of the SINR, and hence the sum rate, for the BF and RZF precoders with perfect CSI estimation and common correlation for all the users. This simplified analysis for perfect CSI estimation is used as a baseline for comparison with imperfect channel estimation (see subsection 4.4.1). Recalling the analytical P-DoF channel model given in (2.6), and due to the condition of common spatial channel correlation considered in this chapter, the covariance matrices of the users are all the same and of the form $\mathbf{R} = \frac{N}{P}\mathbf{A}\mathbf{A}^{\mathrm{H}}$, where $\mathbf{A} \in \mathbb{C}^{N\times P}$ is constructed from $P \leq N$ columns of an arbitrary $N \times N$ unitary matrix so that $\mathbf{A}^{\mathrm{H}}\mathbf{A} = \mathbf{I}_P$. To this end, a closed-form analysis of the SINR, and hence the sum rate, is developed when the BS employs BF precoding with perfect CSI estimation and the P-DoF model is used. Specifically, substituting $\mathbf{R}$ of the P-DoF model into (4.3) yields, after some algebraic manipulations, the asymptotic tight approximations of the SINR for the BF precoder with perfect CSI estimation, which is summarised in the following Theorem.

Theorem 4.1. *An asymptotically tight approximation of the SINR for BF precoding under perfect channel estimation and based on the P-DoF channel model with common spatial correlation is given as*

$$\overline{\mathrm{SINR}}^{\mathrm{BF,\,PF}} = \frac{N}{\frac{1}{\rho_{\mathrm{d}}} + \frac{KN}{P}}. \tag{4.14}$$

For the high SNR case, so that when $\rho_{\mathrm{d}} \to \infty$, an analytical closed-form solution for the achievable sum rate with perfect channel estimation emerges. In this special case the achievable sum rate $\bar{\mathrm{C}}^{\mathrm{BF,\,PF},\infty}$, further simplifies to

$$\bar{\mathrm{C}}^{\mathrm{BF,PF},\infty} = K \cdot \log_2\left(1 + \frac{P}{K}\right), \tag{4.15}$$

where P represents the number of degrees of freedom in the physical channel model.

Proof: A proof of Theorem 4.1 is provided in Appendix B. ■

The expression in (4.15) shows explicitly the dependence of the achievable sum rate performance on the number of degrees of freedom that the propagation channel can offer as well as the number of served users under a common spatial correlation. Specifically, the growing in sum rate performance of the BF precoder under perfect channel estimation is limited by the number of users and number of degrees of freedom, which implicitly depends on N. The following Theorem provides a simplified expression for the achievable sum rate when the BS employs RZF precoding with perfect channel estimation and the P-DoF channel model is utilised.

Theorem 4.2. *An asymptotically tight approximation of the SINR for the RZF precoding uder perfect channel estimation and based on the P-DoF channel model with common spatial correlation is given as*

$$\overline{\mathrm{SINR}}^{\mathrm{RZF,PF}} = \frac{N\mathring{\xi}\,\mathring{\delta}^2}{\frac{(1+\mathring{\delta})^2}{\rho_{\mathrm{d}}} + K\mathring{\xi}\,\mathring{\mu}}, \tag{4.16}$$

where $\mathring{\delta}$ is a closed-form solution to the fixed-point equation provided in (4.5)–(4.6) *when common spatial correlation between users is considered and reads*

$$\mathring{\delta} = \frac{P - K - \mathring{Z} + \sqrt{(K-P)^2 + 2(K+P)\mathring{Z} + \mathring{Z}^2}}{2\mathring{Z}} \tag{4.17}$$

where parameter $\mathring{Z}$ is given as $\mathring{Z} = P\zeta$, where ζ is the regularisation parameter for the RZF precoder. Parameters $\mathring{\xi}$ and $\mathring{\mu}$ are simplified versions of (4.10) *and* (4.11) *and defined respectively as*

$$\mathring{\xi} = \frac{\mathring{Q}\,\zeta\left(\left(1+\mathring{\delta}\right)^2\mathring{Z}^2 + K\left(2\mathring{Q} - K - P\right)\right)}{\left(1+\mathring{\delta}\right)P\,Z\left(\mathring{Q} - P\right)} \tag{4.18}$$

$$\mathring{\mu} = \frac{(1+\mathring{\delta})^2 N P\Big(\big(1+\mathring{\delta}\big)^2 \mathring{Z}^3 - 2P^2 K\zeta - K(K-2\mathring{Q})\mathring{Z} + P^2\zeta\big(P - 2\big(1+\mathring{\delta}\big)\mathring{Z}\big)\Big)}{\mathring{Q}^2 P\zeta\left(\big(1+\mathring{\delta}\big)^2 \mathring{Z}^2 + K\big(2\mathring{Q} - K - P\big)\right)} \tag{4.19}$$

where $\mathring{Q} = K + \mathring{Z} + \mathring{\delta}\mathring{Z}$ *is used for notational convenience. The asymptotic tight analysis for the achievable sum rate of the RZF precoding is obtained by substituting* (4.16) *into* (4.2)*.*

Appendix D explains how the SINR expression in (4.16) for the RZF precoder with perfect CSI estimation and common spatial correlation between users is obtained. The SINR expression for the RZF precoder with perfect CSI estimation is simplified, which depends only on P, K, N, and ρ_{d}. Thus, the sum rate of RZF precoding for the massive MIMO system under consideration can be readily obtained when these relevant system parameters are known, which obviates the need for computationally intensive Monte Carlo simulations. In Chapter 5, the simplified SINR expression for the BS and RZF precoders developed in this section will be applied to a more general scenario, where users exhibit different spatially correlated channels. The following section investigates the problem of CSI estimation using DL training sequences in an FDD massive MIMO communication system.

4.3 Achievable sum rate analysis and training sequence optimisation in FDD massive MIMO systems with common correlation between users and imperfect channel estimation

This section aims to establish a tractable framework analysis that can characterise the achievable sum rate performance of the BF and RZF precoders with imperfect CSI estimation when a DL training sequence is considered. To this end, a DL system model, which is introduced in Chapter 3 section 3.2, is used. As described in Chapter 3, during the training transmission phase, the BS transmits training sequences of length T_{p} symbols, while the remaining symbols

correspond to the data transmission phase.

Due to the imperfect channel estimation considered in this section, the BS employs precoding using the estimate $\hat{\mathbf{H}}$ of the DL channel $\mathbf{H}$. In particular, the BF and RZF precoders now depend on the imperfect channel estimation through $\hat{\mathbf{H}}$. Accordingly, the average achievable sum rate, C, for an FDD massive MIMO system with imperfect channel estimation can be expressed as

$$C = \left(1 - \frac{T_{\mathrm{p}}}{T_{\mathrm{c}}}\right) \cdot K \cdot \log_2\big(1 + \mathrm{SINR}\big), \tag{4.20}$$

where $T_{\mathrm{c}} \in \mathbb{Z}^+$ is the coherence time in symbols and SINR term is the effective SINR expression given in (3.5) with imperfect CSI estimation and statistically equal correlation for all the users.

Recalling the MMSE channel estimate introduced in Chapter 3 section 3.3, and based on the condition of a common spatial channel correlation for all users, the covariance matrix of the MMSE channel estimation $\mathbf{\Phi} \in \mathbb{C}^{N \times N}$ in (3.10) can be rewritten as in (4.21)

$$\mathbf{\Phi} = \sqrt{\rho_{\mathrm{p}}}\mathbf{R}\mathbf{S}_{\mathrm{p}}\big(\rho_{\mathrm{p}}\mathbf{S}_{\mathrm{p}}^{\mathrm{H}}\mathbf{R}\mathbf{S}_{\mathrm{p}} + \mathbf{I}_{T_{\mathrm{p}}}\big)^{-1}\sqrt{\rho_{\mathrm{p}}}\mathbf{S}_{\mathrm{p}}^{\mathrm{H}}\mathbf{R}. \tag{4.21}$$

As explained in Chapter 2, minimising the mean square error of the channel estimate for a given ρ_{p} maximises the $\log_2(1 + \mathrm{SINR})$ term but it comes at the cost of the pre-log fraction term $(1 - T_{\mathrm{p}}/T_{\mathrm{c}})$. To maximise the sum rate of an FDD massive MIMO system for arbitrary values of N and T_{c}, both terms should be taken into account. This chapter focuses on finding a non-iterative solution for the pilot structure $\mathbf{S}_{\mathrm{p}}$ based on an objective to maximise the achievable sum rate of the massive MIMO system under consideration with limited coherence time. This chapter also aims to develop a tractable closed-form solution for the optimum training duration that maximises the achievable sum rate in an FDD massive MIMO system. In order to achieve the aforementioned fundamental objectives, this chapter starts with a simplified scenario by considering that all the users exhibit the same spatial correlation. Consequently, the MMSE channel estimates $\mathbf{\Phi}$ of all the users become statistically the same, which makes the problem of CSI estimation to be analytically tractable. In particular, the reason for considering such a simplified scenario is to explore the optimum training sequence design, which can be used for all the users. This would also allow the optimum training sequence length to be mathematically derived and expressed analytically in closed-form (see subsection 4.3.3). Importantly, this

tractable framework analysis will be extended in Chapter 5 to provide a low-complexity feasible solution for the DL training design in a more general scenario of the FDD massive MIMO systems where users exhibit different spatial correlations.

Using the EVD described in Chapter 2 subsection 2.5.1, the channel covariance matrix $\mathbf{R}$ at the BS with a common spatial correlation between users can be decomposed into

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}}, \tag{4.22}$$

where $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_N] \in \mathbb{C}^{N \times N}$ is a unitary matrix of the eigenvectors of $\mathbf{R}$ and $\mathbf{\Lambda} \in \mathbb{R}^{N \times N}$ is a diagonal matrix of the eigenvalues of $\mathbf{R}$ arranged in descending order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$. The pilot sequence for the DL training transmission phase is designed based on eigenvectors of the channel covariance matrix. In particular, the spatio-temporal training sequence matrix $\mathbf{S}_{\mathrm{p}} \in \mathbb{C}^{N \times T_{\mathrm{p}}}$ is constructed from the first T_{p} eigenvectors of the channel covariance matrix $\mathbf{R}$, corresponding to the largest eigenvalues, as given in (4.23).

$$\mathbf{S}_{\mathrm{p}} = \mathbf{U}_{T_{\mathrm{p}}} = \left[\mathbf{u}_1, \ldots, \mathbf{u}_{T_{\mathrm{p}}}\right] \tag{4.23}$$

Substituting (4.23) into (4.21) simplifies the covariance matrix of the MMSE channel estimate as

$$\mathbf{\Phi} = \rho_{\mathrm{p}} \mathbf{U}_{T_{\mathrm{p}}} \mathbf{\Lambda}_{T_{\mathrm{p}}}^2 \left(\rho_{\mathrm{p}} \mathbf{\Lambda}_{T_{\mathrm{p}}} + \mathbf{I}_{T_{\mathrm{p}}}\right)^{-1} \mathbf{U}_{T_{\mathrm{p}}}^{\mathrm{H}}, \tag{4.24}$$

where $\mathbf{\Lambda}_{T_{\mathrm{p}}} \in \mathbb{R}^{T_{\mathrm{p}} \times T_{\mathrm{p}}}$ is a diagonal matrix with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{T_{\mathrm{p}}}$. In principle, increasing T_{p} allows for more training signal energy to be received, but it comes at the cost of reduced achievable sum rate due to a shorter data transmission. This implies that the energy in the channel that corresponds to the last eigenvectors $\mathbf{u}_{T_{\mathrm{p}}+1}, \ldots, \mathbf{u}_N$ is not used in precoding.

4.3.1 Formulation of the achievable sum rate optimisation problem in FDD with imperfect channel estimation

Maximising the DL achievable sum rate over the training sequence duration in an FDD massive MIMO system under consideration equates to the optimisation problem defined in (4.25).

$$\begin{aligned} \underset{T_\mathrm{p}}{\text{maximise}} \quad & \left(1-\frac{T_\mathrm{p}}{T_\mathrm{c}}\right)\cdot K\cdot \log_2\big(1+\mathrm{SINR}\big) \\ \text{subject to} \quad & 1\le T_\mathrm{p}\le \min\big(N,T_\mathrm{c}\big) \end{aligned} \tag{4.25}$$

The optimisation problem above corresponds to a K user massive MIMO system where each user has the same spatial correlation at the BS and, as a result, equal ergodic achievable rates. Though the problem setup may seem straightforward, solving the problem formulated in (4.25) for arbitrary correlation matrix $\mathbf{R}$ and different values of N is computationally demanding since the expectations in the SINR expression in (3.5) need to be obtained for different choices of T_p using extensively Monte Carlo simulations. A computationally feasible solution for finding an optimum training sequence length T_p^* may be obtained by invoking a large system limit based on random matrix theory methods [30]. In particular, using a random matrix theory allows the SINR term in (3.5) for the BF and RZF precoders with imperfect channel estimation to be replaced with the approximations developed in this subsection, so that the revised optimisation problem can be expressed as in (4.26).

$$\begin{aligned} \underset{T_\mathrm{p}}{\text{maximise}} \quad & \left(1-\frac{T_\mathrm{p}}{T_\mathrm{c}}\right)\cdot K\cdot \log_2\big(1+\overline{\mathrm{SINR}}\big) \\ \text{subject to} \quad & 1\le T_\mathrm{p}\le \min\big(N,T_\mathrm{c}\big) \end{aligned} \tag{4.26}$$

where $\overline{\mathrm{SINR}}$ is the asymptotic approximation for the effective SINR term in (3.5) with an imperfect channel estimation and common correlation for all the users.

The following Propositions (i.e., Propositions 4.3 and 4.4) elaborate the analytical results of the BF and RZF precoders, respectively, for a general channel correlation model based on the random matrix theory methods [30]. It is worth noting that Propositions 4.3 and 4.4 are modified versions of Theorem 4 and Theorem 6 in [30], for the massive MIMO system model under consideration, where the imperfect channel estimation is carried out on the DL of an FDD massive MIMO with signal-cell communication system. Furthermore, Propositions 4.3 and 4.4 consider a special scenario where all users exhibit the same spatial correlation at the BS.

Proposition 4.3. *Let* $\mathrm{SINR}^{\mathrm{BF}}$ *denote the SINR for BF precoding as given in* (3.5)*. An asymptotically tight approximation of* $\mathrm{SINR}^{\mathrm{BF}}$ *of an FDD massive MIMO systems with common spatial*

correlation between users and imperfect channel estimation is given as

$$\overline{\mathrm{SINR}}^{\mathrm{BF}} = \frac{\left(\mathrm{tr}(\mathbf{\Phi})\right)^2}{\frac{1}{\rho_\mathrm{d}}\mathrm{tr}(\mathbf{\Phi}) + K\,\mathrm{tr}(\mathbf{R}\mathbf{\Phi})}, \tag{4.27}$$

where $\mathbf{\Phi}$ *is the covariance matrix of the MMSE channel estimate under a common correlation between users, which is provided in* (4.21)*.*

The SINR approximation given in (4.27) is generally valid for any channel correlation model and training sequence type. For the training sequence matrix in (4.23) that is common for all the users, it is straightforward to obtain simplified expressions for the covariance matrix of the MMSE channel estimate $\mathrm{tr}(\mathbf{\Phi})$ and $\mathrm{tr}(\mathbf{R}\mathbf{\Phi})$ from (4.24) as

$$\mathrm{tr}(\mathbf{\Phi}) = \sum_{n=1}^{T_\mathrm{p}} \frac{\lambda_n^2}{\lambda_n + 1/\rho_\mathrm{p}}, \tag{4.28}$$

$$\mathrm{tr}(\mathbf{R}\mathbf{\Phi}) = \sum_{n=1}^{T_\mathrm{p}} \frac{\lambda_n^3}{\lambda_n + 1/\rho_\mathrm{p}}, \tag{4.29}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{T_\mathrm{p}}$ are the $T_\mathrm{p} \leq N$ largest eigenvalues of the channel covariance matrix $\mathbf{R}$. Appendix E shows how the expressions in (4.28) and (4.29) are obtained.

Proposition 4.4. *Let* $\mathrm{SINR}^{\mathrm{RZF}}$ *denote the* SINR *of RZF precoding as defined in* (3.5)*. An asymptotically tight approximation for the* $\mathrm{SINR}^{\mathrm{RZF}}$ *of an FDD massive MIMO system with common spatial correlation between users and imperfect channel estimation is provided as*

$$\overline{\mathrm{SINR}}^{\mathrm{RZF}} = \frac{N\,\bar{\xi}\,\delta^2}{\frac{(1+\delta)^2}{\rho_\mathrm{d} K} + \bar{\xi} K \bar{\mu}}, \tag{4.30}$$

where the parameter $\bar{\xi} \in \mathbb{R}$ *is defined later in* (4.36) *and* $\delta \in \mathbb{R}$ *is an auxiliary variable arises from the asymptotic random matrix theory, which can be obtained using a fixed-point algorithm. Defining a recursion on integer* t*, where* $t = 1, 2, \ldots,$

$$\delta^{(t)} = \frac{1}{N}\mathrm{tr}\left(\mathbf{\Phi}\left(\frac{K}{N}\frac{\mathbf{\Phi}}{(1+\delta^{(t-1)})} + \zeta \mathbf{I}_N\right)^{-1}\right), \tag{4.31}$$

with an initial value $\delta^{(0)} = 1/\zeta$*, where* $\zeta > 0$ *is the regularisation parameter of the RZF*

precoder and the variable $\delta \in \mathbb{R}$ *is obtained as the limit*

$$\delta = \lim_{t \to \infty} \delta^{(t)}, \tag{4.32}$$

After the solution of the fixed-point equations in (4.31) *and* (4.32) *is numerically obtained, it is substituted into*

$$\mathbf{T} = \left(\zeta \mathbf{I}_N + \frac{K}{N} \frac{\mathbf{\Phi}}{1+\delta} \right)^{-1}, \tag{4.33}$$

to obtain matrix $\mathbf{T} \in \mathbb{C}^{N \times N}$. *Auxiliary matrix* $\bar{\mathbf{T}} \in \mathbb{C}^{N \times N}$ *is given by*

$$\bar{\mathbf{T}} = \mathbf{T} \left(\mathbf{I}_N + \frac{K}{N} \frac{\mathbf{\Phi} \bar{\delta}}{(1+\delta)^2} \right) \mathbf{T}, \tag{4.34}$$

and the auxiliary variable $\bar{\delta} \in \mathbb{R}$ *is obtained as*

$$\bar{\delta} = \left(1 - \frac{\mathrm{tr}\big(\mathbf{\Phi T \Phi T}\big)}{\big(N(1+\delta)\big)^2} \right)^{-1} \frac{1}{N} \mathrm{tr}\big(\mathbf{T \Phi T}\big), \tag{4.35}$$

Parameter $\bar{\xi} \in \mathbb{R}$ *is given by substituting* $\mathbf{T}$ *and* $\bar{\mathbf{T}}$ *into*

$$\bar{\xi} = \big(\mathrm{tr}(\mathbf{T}) - \zeta \mathrm{tr}(\bar{\mathbf{T}}) \big)^{-1}. \tag{4.36}$$

The auxiliary variable $\bar{\mu} \in \mathbb{R}$ *is determined based on the dominated convergence and the continuous mapping theorems as developed in [30], which is adopted here to be applicable for the massive MIMO system under consideration, and thus, it is obtained from the expressions given in* (4.37)*,* (4.38) *and* (4.39)

$$\bar{\mu} = \frac{1}{N} \mathrm{tr}\big(\mathbf{R T'}\big) - \frac{2\mathrm{Re}\big(\mathrm{tr}\big(\mathbf{\Phi T}\big)\mathrm{tr}\big(\mathbf{\Phi T'}\big)\big)(1+\delta) - \mathrm{tr}\big(\mathbf{\Phi T}\big)^2 \delta'}{\big(N(1+\delta)\big)^2} \tag{4.37}$$

$$\mathbf{T'} = \mathbf{T} \left(\mathbf{\Phi} + \frac{K}{N} \frac{\mathbf{\Phi} \delta'}{(1+\delta)^2} \right) \mathbf{T} \tag{4.38}$$

$$\delta' = \left(1 - \frac{\mathrm{tr}\big(\mathbf{\Phi T \Phi T}\big)}{\big(N(1+\delta)\big)^2} \right)^{-1} \frac{1}{N} \mathrm{tr}\big(\mathbf{T \Phi T \Phi}\big) \tag{4.39}$$

The SINR analysis for the RZF precoder with imperfect channel estimation provided in (4.30) is generally valid for any channel correlation model and training sequence type. This analysis is given in terms of several auxiliary variables that arise from the asymptotic random matrix theory methods [30]. A special case where the performance of the RZF precoder with imperfect channel estimation can be characterised in closed-form is presented in subsection 4.3.4 using the P-DoF channel model. The following subsection investigates an analytical closed-form solution for the optimum training sequence length of BF precoding for the system under consideration in the P-DoF channel model.

4.3.2 Sum rate analysis of the BF precoder with imperfect channel estimation based on the P-DoF channel model and common correlation between users in FDD systems

Although the SINR expression in Proposition 4.3 can be further simplified for a particular system under consideration by substituting (4.28) and (4.29) for $\mathbf{R}$ and $\mathbf{\Phi}$, respectively, determining the sum rate for the BF precoder in a form that is useful for the optimisation problem defined in (4.26) is still challenging. In order to solve this and gain further insight into the problem, the analytical P-DoF channel model, as defiened in Chapter 2 subsection 2.7.1, is considered in this subsection. This supports the formulation of both $\overline{\text{SINR}}^{\text{BF}}$ and $\overline{\text{SINR}}^{\text{RZF}}$, leading to the successful computation of their respective sum rates with very low computational complexity. Recalling the P-DoF channel model defined in (2.6), and due to the condition of common correlation for all the users, the covariance matrices of all the users are the same and of the form $\mathbf{R}_k = \mathbf{R} = \frac{N}{P}\mathbf{A}\mathbf{A}^{\text{H}}\,\forall k$, where $\mathbf{A} \in \mathbb{C}^{N\times P}$ is constructed from $P \leq N$ columns of an arbitrary $N \times N$ unitary matrix so that $\mathbf{A}^{\text{H}}\mathbf{A} = \mathbf{I}_P$. This implies that the EVD of $\mathbf{R}$, defined in (2.4), where $\mathbf{\Lambda}$ is a diagonal matrix containing the ordered eigenvalues of $\mathbf{R}$, namely, $\lambda_1 = \lambda_2 = \cdots = \lambda_P = N/P$ and $\lambda_{P+1} = \cdots = \lambda_N = 0$. Substituting the EVD of the P-DoF channel model into (4.24) provides the channel estimate covariance matrix as

$$\mathbf{\Phi} = \frac{\rho_{\text{p}}}{(\rho_{\text{p}} + P/N)\,P/N}\mathbf{U}_m\mathbf{U}_m^{\text{H}}, \quad m = \begin{cases} T_{\text{p}} & \text{if } T_{\text{p}} \leq P; \\ P & \text{otherwise,} \end{cases} \tag{4.40}$$

where $\mathbf{U}_m \in \mathbb{C}^{N\times m}$ denotes a matrix constructed from m eigenvectors of $\mathbf{R}$, as discussed in Section 4.3. Since the channel has $P \leq N$ degrees of freedom and the energy consumed by the training-phase is $\rho_{\mathrm{p}}T_{\mathrm{p}}$, choosing $T_{\mathrm{p}} > P$ while keeping ρ_{p} constant leads to the same channel estimation error performance as $T_{\mathrm{p}} = P$ but unnecessarily consumes more energy in the training phase. Therefore, $T_{\mathrm{p}} \leq P$ is always taken in the following sequel. The above analysis indicates that the constraint in the optimisation problem defined in (4.26) should be revised for the P-DoF channel as $T_{\mathrm{p}} \leq \min(P, T_{\mathrm{c}})$, which limits the training sequence length when the channel is stochastically rank deficient.

To obtain a closed-form analysis for the SINR of the BF precoder based on the P DoF channel model, start by combining (4.40) and the EVD of $\mathbf{R}$, as discussed above, with (4.27)–(4.29). Straightforward algebra provides an SINR approximation for the BF precoder with common spatial correlation between users as

$$\overline{\mathrm{SINR}}^{\mathrm{BF}} = \frac{\rho_{\mathrm{d}}\rho_{\mathrm{p}}T_{\mathrm{p}}}{\big(P/N + \rho_{\mathrm{p}}\big)\big(P/N + K\rho_{\mathrm{d}}\big)}, \qquad T_{\mathrm{p}} \leq P. \tag{4.41}$$

Proof: A proof of expression (4.41) is presented in the Appendix F. ■

Comparing (4.41) to Eq. (28) in [30], which provides an SINR approximation for the BF precoder in TDD massive MIMO when length K uplink pilot sequences are transmitted, shows that the effect of DL channel estimation on SINR is to replace the DoF of the channel P by the training sequence length $T_{\mathrm{p}} \leq P$. With this choice, however, the loss in achievable rate due to training, $(1 - T_{\mathrm{p}}/T_{\mathrm{c}})$, overcomes the increase in SINR when P become sufficiently large. This intuitive statement is formalised for the case of the high SNR region described in Theorem 4.3 in the following subsection. In Chapter 5, the asymptotic analysis in (4.41) is extended to provide a simplified expression for the SINR of the BF precoder in a more general scenario where the users exhibit different spatial correlations.

4.3.3 Training sequence length optimisation for BF precoder with common spatial correlation in FDD

This subsection shows how the optimum DL training sequence length that maximises the achievable sum rate is obtained analytically for the BF precoder at high SNR in the P-DoF

channel model. To the best of the author's knowledge, no analytical solution for the optimum pilot length that maximises the achievable sum rate has yet been proposed in an FDD based massive MIMO due to the technical challenge of deriving a closed-form. This forms a key contribution of this research. Substituting (4.41) into (4.26) provides a fast numerical optimisation method for the BF precoder. Also, for the case of high SNR, $\rho_{\mathrm{d}} = \rho_{\mathrm{p}} = \rho \to \infty$, an analytical solution emerges. In this special case the average achievable sum rate further simplifies to

$$\bar{C}^{\mathrm{BF},\infty} = \left(1 - \frac{T_{\mathrm{p}}}{T_{\mathrm{c}}}\right) \cdot K \cdot \log_2\left(1 + \frac{T_{\mathrm{p}}}{K}\right), \qquad T_{\mathrm{p}} \leq P, \tag{4.42}$$

which leads to the following Theorem. The proof of the achievable sum rate expression in (4.42) is provided in Appendix G.

Theorem 4.3. *For K users with common spatial correlation between users, P-DoF channel model and channel coherence time T_{c}, a novel DL training sequence length T_{p}^{*} that maximises the average achievable sum rate, in an FDD massive MIMO system using BF precoding, at high SNR and with uniform power allocation $\rho_{\mathrm{d}} = \rho_{\mathrm{p}}$, is*

$$T_{\mathrm{p}}^{*} = \begin{cases} P & \text{if } P < \tau, \\ \underset{T_{\mathrm{p}} \in \{\lceil \tau \rceil \lfloor \tau \rfloor\}}{\arg\max}\ \bar{C}^{\mathrm{BF},\infty} & \text{otherwise}, \end{cases} \tag{4.43}$$

where

$$\tau = \frac{K + T_{\mathrm{c}}}{W\left(\frac{K+T_{\mathrm{c}}}{K}\mathrm{e}\right)} - K. \tag{4.44}$$

The Lambert W-function $W(\cdot)$ is defined in [133] and e *is Euler's number. In* (4.43)*, $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ denote the ceiling and floor functions, respectively, which accommodates the necessary integer value of T_{p} given that $\tau \in \mathbb{R}$.*

Proof: A proof of Theorem 4.3 is provided in the Appendix H. ■

From Theorem 4.3, the optimal DL training sequence length for the BF precoder in the P-DoF channel at high SNR is characterised as follows. It is equal to the degrees of freedom of the channel when P is below τ and saturates at, or below, $\lceil \tau \rceil$ when the DoF exceeds (4.44). The achievable sum rate with DL channel estimation and BF precoding can also be upper bounded

using Theorem 4.3. Specifically, given K and T_c, the achievable sum rate for BF precoding with $\rho_d = \rho_p$ is upper bounded by

$$\bar{C}^{\mathrm{BF},\infty} \leq \left(1 - \frac{\tau}{T_c}\right) \cdot K \cdot \log_2\left(1 + \frac{\tau}{K}\right). \tag{4.45}$$

Since τ does not depend on P or N, the rate saturates at a constant level below (4.45) when the DoF exceeds (4.44), i.e. $P > \tau$. This also means that the asymptotic sum rate of the P-DoF channel with BF precoding, as a function of N, is independent of $P/N \in (0, 1]$, although the more rank deficient the correlation matrix is the more BS antennas are needed to approach (4.45). It should be pointed out that this behaviour is in contrast to BF precoding with UL training where the SINR, and hence the sum rate, grows without bound as a function of P when K and T_c are fixed. The following subsection explores an analytical asymptotic approximation for the SINR of RZF precoding with imperfect channel estimation based on the P-DoF channel model and a common spatial correlation for all the users.

4.3.4 Sum rate analysis for the DL RZF precoder based on the P-DoF channel model and common correlation in FDD with imperfect channel estimation

This subsection develops a closed-form analysis for the SINR, and hence the sum rate, when the BS employs RZF precoding and the P-DoF channel model is used with common spatial correlation between users. The following Theorem summarises the key results of the RZF precoder in an FDD massive MIMO system with imperfect channel estimation and when users exhibit the same correlation at the BS.

Theorem 4.4. *An asymptotically tight approximation for the SINR of the RZF precoder in an FDD massive MIMO system based on the P-DoF channel model with common spatial correlation for all the users is given as*

$$\overline{\mathrm{SINR}}^{\mathrm{RZF}} = \frac{N\tilde{\xi}\,\tilde{\delta}^2}{\frac{(1+\tilde{\delta})^2}{\rho_d} + K\tilde{\xi}\,\tilde{\mu}}, \qquad T_p \leq P, \tag{4.46}$$

where $\tilde{\delta}$ is a closed-form solution to the fixed-point equations in (4.31)–(4.32) *when* $\mathbf{R}_k = \mathbf{R}$

and $\mathbf{\Phi}_k = \mathbf{\Phi}\ \forall k = 1, \ldots, K$ *and reads*

$$\tilde{\delta} = \frac{T_\mathrm{p} - K - \tilde{Z} + \sqrt{\left(K - T_\mathrm{p}\right)^2 + 2\left(K + T_\mathrm{p}\right)\tilde{Z} + \tilde{Z}^2}}{2\tilde{Z}}, \tag{4.47}$$

where $\tilde{Z} = P\zeta\left(1 + 1/\rho_\mathrm{p}\right)$. *Parameters* $\tilde{\xi}$ *and* $\tilde{\mu}$ *are simplified versions of* (4.36) *and* (4.37) *with common spatial correlation for all the users, which are obtained from the expressions given in* (4.48) *and* (4.49)*, with* $\tilde{Q} = K + \tilde{Z} + \tilde{\delta}\tilde{Z}$.

$$\tilde{\xi} = \frac{\tilde{Q}\,\zeta\left(\left(1 + \tilde{\delta}\right)^2 \tilde{Z}^2 + K\left(2\tilde{Q} - K - T_\mathrm{p}\right)\right)}{\left(1 + \tilde{\delta}\right) T_\mathrm{p}\, Z\left(\tilde{Q} - T_\mathrm{p}\right)} \tag{4.48}$$

$$\tilde{\mu} = \frac{(1 + \tilde{\delta})^2 N\, T_\mathrm{p}\Big(\left(1 + \tilde{\delta}\right)^2 \tilde{Z}^3 - 2KP\zeta T_\mathrm{p} - K(K - 2\tilde{Q})\tilde{Z} + P\zeta T_\mathrm{p}\left(T_\mathrm{p} - 2\left(1 + \tilde{\delta}\right)\tilde{Z}\right)\Big)}{\tilde{Q}^2\, P\zeta\left(\left(1 + \tilde{\delta}\right)^2 \tilde{Z}^2 + K\left(2\tilde{Q} - K - T_\mathrm{p}\right)\right)} \tag{4.49}$$

Appendix I indicates how the SINR expression of the RZF precoder in (4.46) for a particular system model under consideration is obtained. From Theorem 4.4, the asymptotic tight approximation of the SINR for RZF precoding with imperfect channel estimation depends on the parameters P, N, K ρ_p. This allows the achievable rate of the RZF precoder in an FDD massive MIMO system with common correlation between users to be readily calculated using (4.26) when the relevant system parameters are known without using computationally demanding Monte Carlo simulations, as carried out in Chapter 3. It is also stated here in a simplified analysis so that the numerical results are readily reproducible. Building on this simplified analysis along with the P-DoF channel model, Chapter 5 will develop a closed-form expression for the RZF precoder but with an extension to a more general scenario with K independent channel covariance matrices. The following section presents the numerical results of the analysis for both BF and RZF precoders.

4.4 Numerical results and discussion

In this section, simulation and theoretical results are presented, which investigate the system performance in terms of the achievable sum rate for the BF and RZF precoders in an FDD massive MIMO. The results are provided for the P-DoF channel model when the salient system parameters K, N, P, T_c, and SNR are used. Comparisons between the sum rates of perfect and imperfect channel estimation for both the BF and RZF precoders are provided in order to instigate the impact of training sequence length and channel estimation on the achievable sum rate performance. In addition, in order to corroborate the findings for a more realistic channel model, numerical results are also provided for the scattering OR channel model.

4.4.1 Comparing system performance for perfect and imperfect channel estimation for common spatial correlation between users

Fig. 4.1 plots achievable sum rate versus the number of BS antennas N for the BF and RZF precoders, comparing perfect and imperfect channel estimation methods. Results are presented for the correlated P-DoF channel model with $P/N = 0.1$[1], $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users. Curves for the BF and RZF precoders with imperfect channel estimation are plotted for computational methods as follows: numerical (BF & RZF) based on equations (4.26), (4.41), (4.46) and simulated (BF & RZF) based on equation (4.25) with the associated SINR term under imperfect channel estimation. Curves for the BF and RZF precoders with perfect channel estimation are plotted as follows: numerical (BF & RZF) based on equations (4.2), (4.15), (4.16) and simulated (BF & RZF) based on equation (4.1) with the associated SINR term under perfect channel estimation. These computational methods provide validation between the theoretical and simulated performances, which show excellent agreement throughout.

As can be observed, for both types of precoders, a larger sum rate is obtained with perfect than imperfect channel estimation, as expected. The gap between sum rate performances of perfect and imperfect channel estimation increases with N. For example, with the BF precoder, the sum rate with imperfect channel estimation saturates at 14 bit/s/Hz whereas the sum rate

[1]For the P-DoF channel model, parameter $P/N = 0.1$ implies relatively strong correlation.

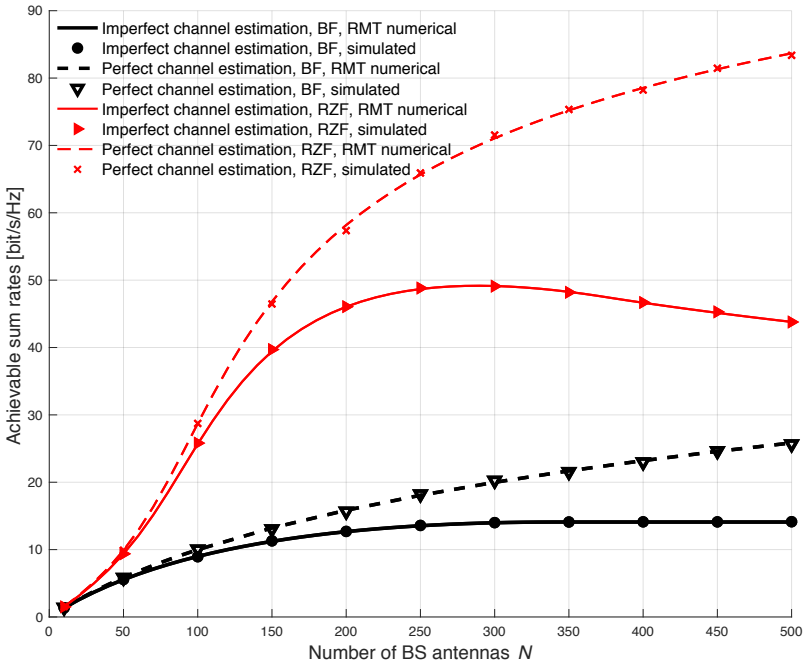


Figure 4.1 Achievable sum rate versus number of BS antennas N based on the P-DoF channel model with $P/N = 0.1$, $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users, comparing DL perfect and imperfect channel estimation.

under perfect channel estimation continues to rise, reaching ~26 bit/s/Hz at $N = 500$. With the RZF precoder, the sum rate difference is greater with the perfect channel estimation, which continues to increase with increasing N and reaches a peak value of ~84 bit/s/Hz at $N = 500$ whereas with imperfect channel estimation, the sum rate peaks at ~49 bit/s/Hz for $N = 294$.

4.4.2 Characterisation of achievable sum rate and training sequence length for common spatial correlation between users

Fig. 4.2 and Fig. 4.3 show plots of the achievable sum rate and the optimum training sequence length T_p^*, respectively, versus the number of BS antennas N, comparing precoder performance in correlated and uncorrelated channels when the P-DoF channel model is used with $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users. Curves for the BF and RZF precoders are plotted for three computational methods as follows: numerical (BF & RZF) based on equations (4.26), (4.41), (4.46); analytical (BF only) based on equations (4.42), (4.43), (4.44); and simulated (BF & RZF) based on equation (4.25) with imperfect channel estimation. These computational methods provide validation between the theoretical and simulated performances,

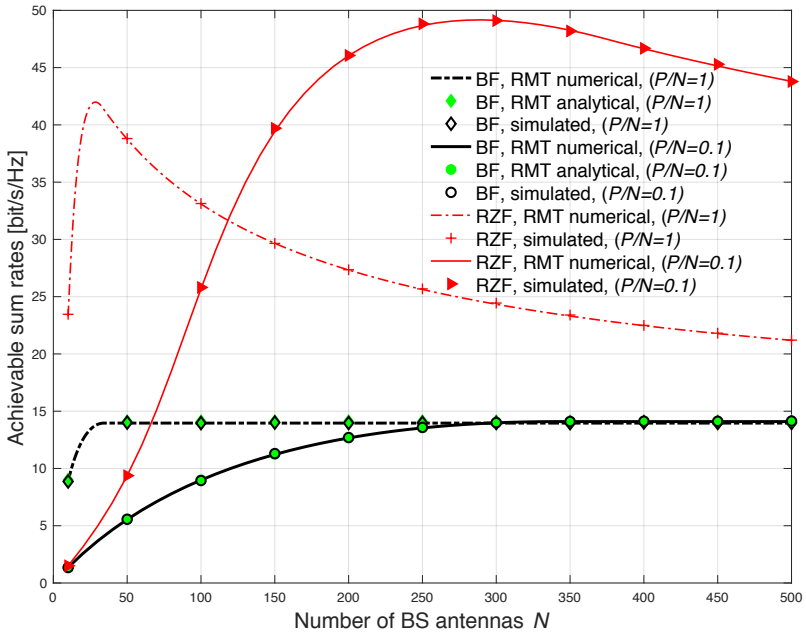


Figure 4.2 Achievable sum rate versus number of BS antennas N for DL channel estimation using the P-DoF channel model with $P/N = \{0.1, 1\}$, $T_c = 100$ symbols, SNR $= 10$ dB and $K = 10$ users.

which show excellent agreement throughout.

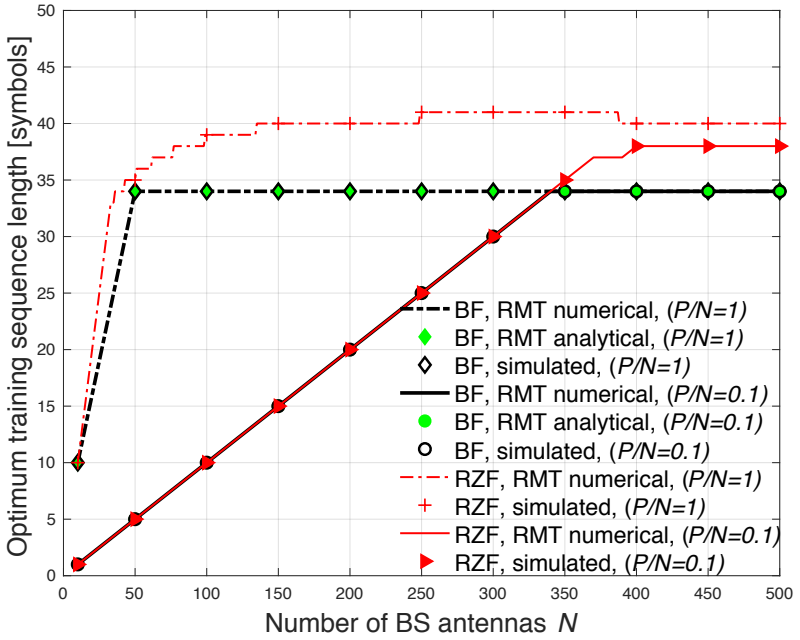


Figure 4.3 Optimum training sequence length T_p^* versus number of BS antennas N for DL channel estimation using the P-DoF channel model with $P/N = \{0.1, 1\}$, $T_c = 100$ symbols, SNR $= 10$ dB and $K = 10$ users.

In Fig. 4.2, the achievable sum rate of the BF precoder, for the uncorrelated channel (i.e. $P/N = 1$), increases steeply with N before saturating at about 14 bit/s/Hz for values of $N > 30$. In contrast, the achievable sum rate for the BF precoder in the correlated channel with $P/N = 0.1$ increases more gradually before also saturating at ~14 bit/s/Hz, though for values of $N > 300$. The saturation of the sum rate for the BF precoder, regardless of the level of correlation in the channel, follows the behaviour predicted by (4.45). For the RZF precoder in the uncorrelated channel, the sum rate gain increases rapidly reaching a peak value of ~41 bit/s/Hz at $N = 29$. For values of $N > 29$, the sum rate slowly decreases monotonically, plateauing at ~21 bit/s/Hz for $N = 500$. However, the sum rate for the RZF precoder in the correlated channel with $P/N = 0.1$ increases more gradually before reaching a larger observed value of ~49 bit/s/Hz at $N = 294$. For values of $N > 294$, the sum rate slightly decreases monotonically reaching ~44 bit/s/Hz at $N = 500$.

In Fig. 4.3, for the uncorrelated channel, where $P/N = 1$[2], the optimum training sequence length T_{p}^{*} initially increases linearly with N for both the BF and RZF precoders. This region of the plots corresponds to when $P < \tau$ in (4.43). Conversely, when $P > \tau$ defines a region where T_{p}^{*} saturates. For the BF precoder, T_{p}^{*} saturates at 34 symbols when $N = 34$ whereas for the RZF precoder T_{p}^{*} saturates at 33 when $N = 33$ but continues to increase gradually for $N > 33$. For the correlated channel, where $P/N = 0.1$, a similar linear characteristic is observed for N up to 340 for the BF precoder and 370 for the RZF precoder. After these regions, T_{p}^{*} saturates at 34 symbols for the BF precoder while T_{p}^{*} continues to increase slightly for the RZF precoder. Importantly, the results in Fig. 4.2 and Fig. 4.3 show that a feasible sum rate can be realised even with the uncorrelated channels, i.e., $P/N = 1$. This is can be justified by the fact that the achievable sum rate is obtained by optimising the training sequence length through the maximisation of $(1 - T_{\mathrm{p}}/T_{\mathrm{c}}) \cdot K \cdot \log_2(1 + \mathrm{SINR})$, instead of only minimising the mean square error of the channel estimate, as typically considered in the conventional analyses of FDD and TDD systems. Furthermore, the results in Fig. 4.2 and Fig. 4.3 confirm that excellent agreement between the numerical, analytical and simulated results was obtained, which forms the fundamental contribution of this research.

[2]For the P-DoF channel model, parameter $P/N = 1$ implies the channels are uncorrelated.

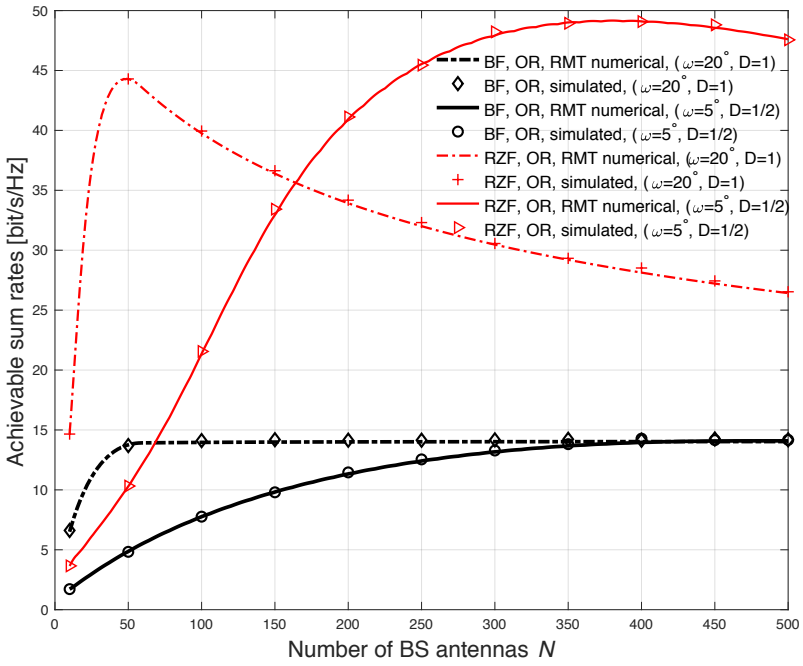


Figure 4.4 Achievable sum rate versus number of BS antennas N with DL channel estimation for $T_c = 100$ symbols, SNR $= 10$ dB, and $K = 10$ users, under the OR channel model with $\theta_H = \pi/6$, $\omega_H = \{5°, 20°\}$, and $D = \{1/2, 1\}$.

4.4.3 Achievable sum rate performance evaluation for the OR channel model

The results in the Subsections 4.4.1 and 4.4.2 were presented based on the analytical P-DoF channel model. In order to validate the applicability of the P-DoF channel model, this subsection considers the application of the OR scattering channel model. Curves for the BF and RZF precoders are plotted as follows: numerical (BF & RZF) based on equations (4.26), (4.27), (4.30) and simulated (BF & RZF) based on equation (4.25) with the associated SINR term based on imperfect CSI estimation.

Fig. 4.4 plots achievable sum rate versus number of BS antennas N for the BF and RZF precoders comparing different levels of correlation over the OR scattering channel model. The parameters for the OR model are chosen with AoA in horizontal (azimuth) direction $\theta_H = 30°$, angular spread in horizontal (azimuth) direction $\omega_H = \{5°, 20°\}$ and normalised antenna spacing $D = \{1/2, 1\}$, where the same system parameters were considered previously in [42, 53]. Note that for the OR model, parameters $\omega_H = 5°$, $D = 1/2$ correspond to strong correlation

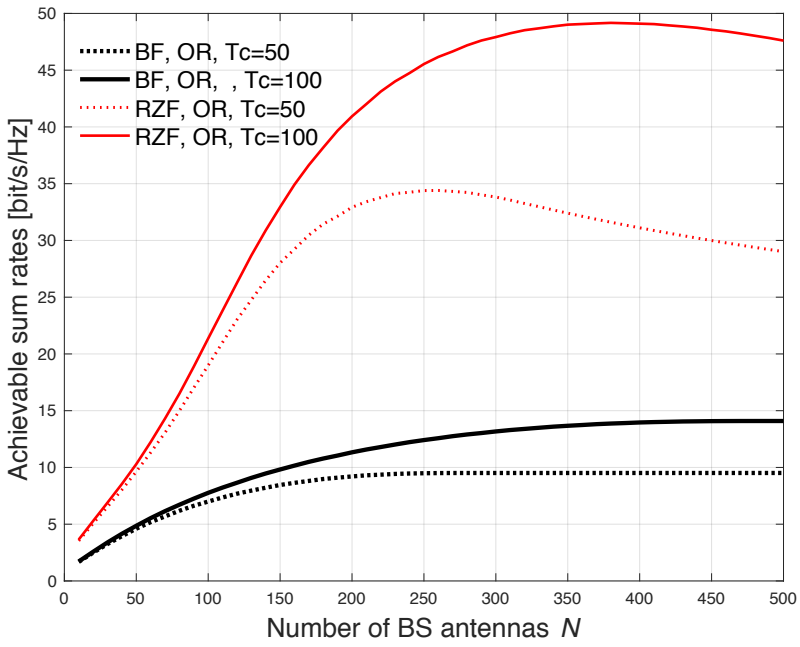


Figure 4.5 Achievable sum rate versus number of BS antennas N with DL channel estimation for $T_c = \{50, 100\}$ symbols, SNR = 10 dB, and $K = 10$ users, under the OR channel model with $\theta_H = \pi/6$, $\omega_H = 5°$, and $D = 1/2$.

whereas $\omega_H = 20°$, $D = 1$ imply relatively weak correlation. Results are shown for the BF and RZF precoders for $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users. The solid lines depict numerical analysis based on random matrix theory, while the coloured markers (red and black) denote simulation. The results demonstrate that excellent agreement between the theoretical and simulated performances are also obtained with the OR channel model. As expected, results in Fig. 4.4 show that, for a massive MIMO system under consideration, the RZF precoder achieves greater sum rate than the BF precoder, both in the correlated and the uncorrelated channels. Also, the results in Fig. 4.4 demonstrate that significant improvement in the sum rate performance is obtained for the RZF precoder when the channels are strongly correlated, i.e. $\omega_H = 5°$, $D = 1/2$. However, the achievable sum rate performance for the BF precoder with weak channel correlation, i.e. $\omega_H = 20°$, $D = 1$, achieves considerable improvement in rate performance over the high correlated channel, i.e. $\omega_H = 5°$, $D = 1/2$ for low N.

Fig. 4.5 plots achievable sum rate versus number of BS antennas N for the BF and RZF precoders comparing different choices of channel coherence time, i.e., $T_c = \{50, 100\}$ symbols, over the OR scattering channel model. Curves for the BF and RZF precoders are obtained

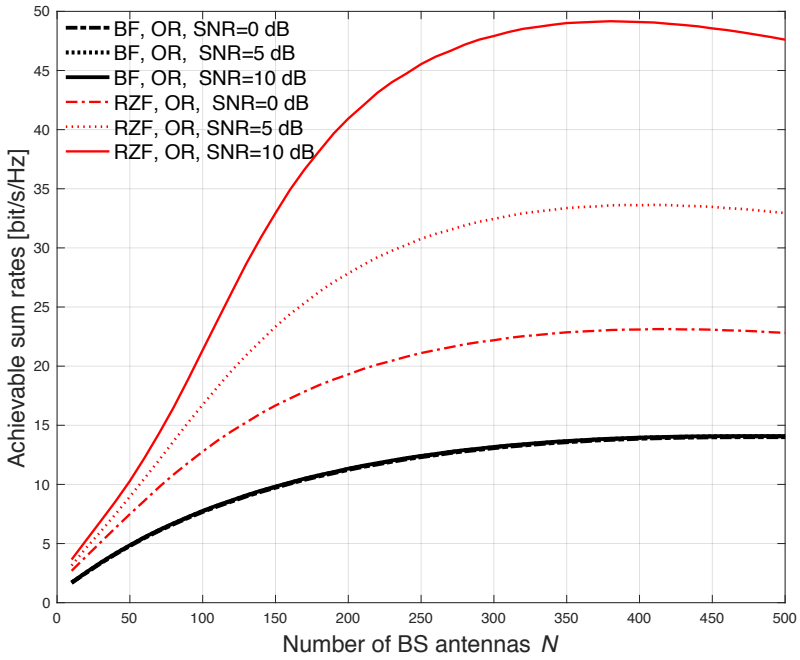


Figure 4.6 Achievable sum rate versus number of BS antennas N with DL channel estimation for $T_c = 100$ symbols, SNR= $\{0, 5, 10\}$ dB, and $K = 10$ users, under the OR channel model with $\theta_H = \pi/6$, $\omega_H = 5°$, and $D = 1/2$.

numerically based on equation (4.26), (4.27), (4.30) with imperfect CSI estimation. The other system parameters are $\theta_H = \pi/6$, $\omega_H = 5°$, $D = 1/2$, SNR = 10 dB and $K = 10$ users. The results in Fig. 4.5 confirm that a considerable improvement in the sum rate performance is obtained for both the BF and RZF precoders when channel coherence time is increased. This is due to the fact that increasing the coherence time results in decreasing the pre-log fraction $(1 - T_p/T_c)$, and thus, increasing the available time for transmitting useful data to the users.

Fig. 4.6 plots achievable sum rate versus number of BS antennas N for the BF and RZF precoders comparing different values of SNR, i.e., SNR=$\{0, 5, 10\}$ dB, over the OR scattering channel model. The other system parameters are $\theta_H = \pi/6$, $\omega_H = 5°$, $D = 1/2$, $T_c = 100$ symbols, and $K = 10$ users. As expected, the results in Fig. 4.6 show that a significant improvement in the sum rate performance is obtained for the RZF precoder when the SNR value is increased. However, for the BF precoder, almost the same rate performances are achieved for the different values of SNR considered.

Fig. 4.7 plots achievable sum rate versus number of BS antennas N for the BF and RZF precoders comparing different numbers of users, i.e., K=$\{6, 8, 12\}$, over the OR scattering channel

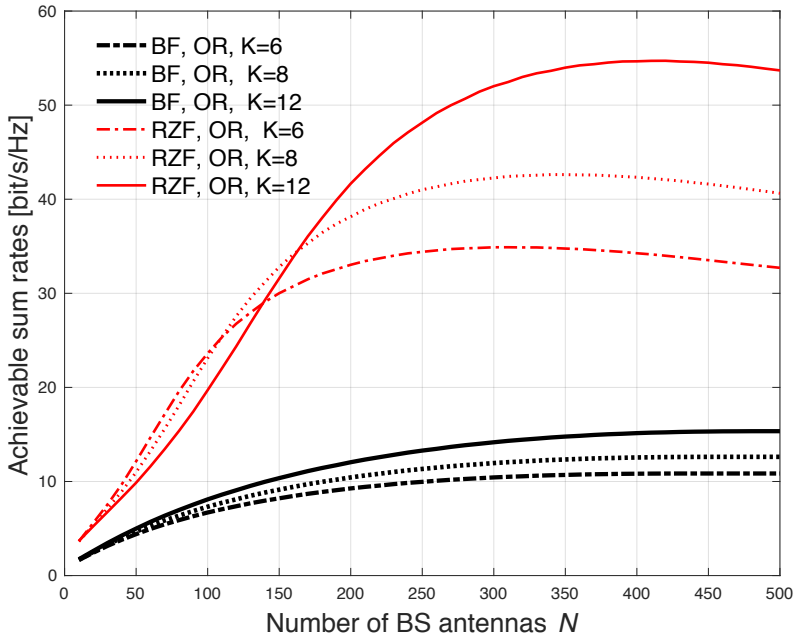


Figure 4.7 Achievable sum rate versus number of BS antennas N with DL channel estimation for $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and K=$\{6, 8, 12\}$ users, under the OR channel model with $\theta_{\mathrm{H}} = \pi/6$, $\omega_{\mathrm{H}} = 5°$, and $D = 1/2$.

model. The other system parameters are $\theta_{\mathrm{H}} = \pi/6$, $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, SNR = 10 dB, and $T_{\mathrm{c}} = 100$ symbols. These curves show the effectiveness of increasing the number of users on the achievable sum rate performance. The results demonstrate that the BF and RZF precoders achieve a considerable improvement in the maximum value of sum rate when the number of users K is increased.

4.4.4 Computational complexity analysis of training sequence design with common correlation between users

In this subsection, the computational complexity analysis of the training sequence optimisation with common correlation between users is provided and compared with the state-of-the-art SCMI/SMSE iterative algorithms for sequence optimisation, which have been already presented in Chapter 3. The overall computational complexity is determined by multiplying the number of iterations each algorithm needs to converge by the number of complex flops involved per iteration. The proposed training design requires only *one iteration*. Table 4.1 gives the complexity analysis in flops by counting the number of multiplications and additions [123] for the

Table 4.1 Computational complexity analysis of the training design with common correlation between users

Algorithm	Complexity in flops
Training sequence with $\mathbf{R}_k = \mathbf{R} \forall k = 1, \ldots, K$	$N^2 r + NT_{\mathrm{p}}$

training sequence optimisation under a common correlation between users. The computational complexity analysis of the SMSE and SCMI sequence designs has been already provided in Chapter 3. Below are the details to describe how the computational complexity of the sequence design when each user has the same spatial correlation was obtained in Table 4.1.

- Determining the EVD of an $N \times N$ rank deficient matrix in (4.22) needs $N^2 r$ flops.
- The scalar matrix multiplication with an $N \times T_{\mathrm{p}}$ matrix requires NT_{p} flops.
- Combining the flops obtained in the above, yields to the complexity analysis in Table 4.1.

Fig. 4.8 and Fig. 4.9 show plots of the computational complexity versus the number of BS antennas N, comparing the training sequence design under a common correlation between users with the SMSE and SCMI training sequence designs. The results in Fig. 4.8 were obtained for the OR model with $\omega_{\mathrm{H}} = 5^{\circ}$, $D = 1/2$ using the pilot sequence length for the BF and RZF precoders, respectively. The other system parameters are given as $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 6$ users. The results in Fig. 4.8 and Fig. 4.9 show that, even though a reduced number of users is used, i.e., $K = 6$, more than four orders-of-magnitude reduction in computational complexity is achieved using the training sequence optimisation based on a common spatial correlation.

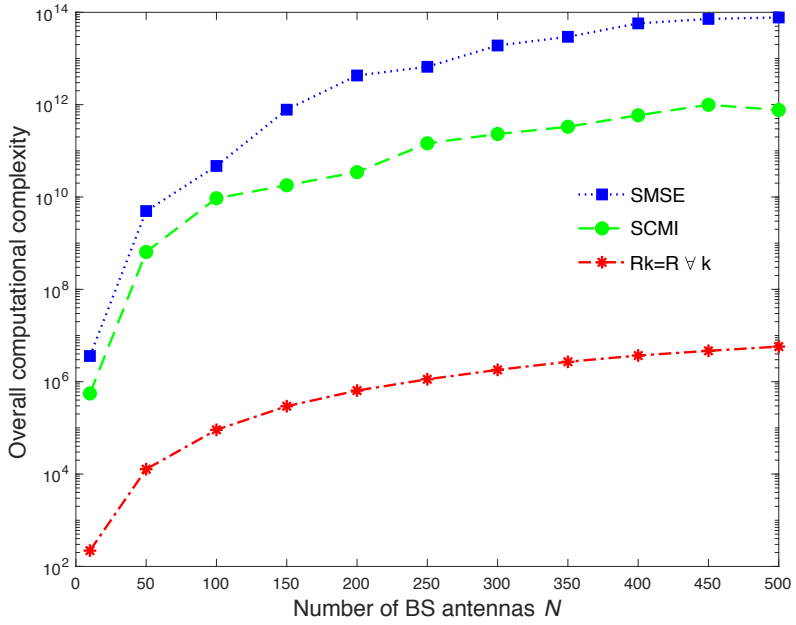


Figure 4.8 Overall computational complexity versus the number of BS antennas N comparing different training sequence methods corresponding to the optimal pilot length in BF precoding in the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$ and $K = 6$.

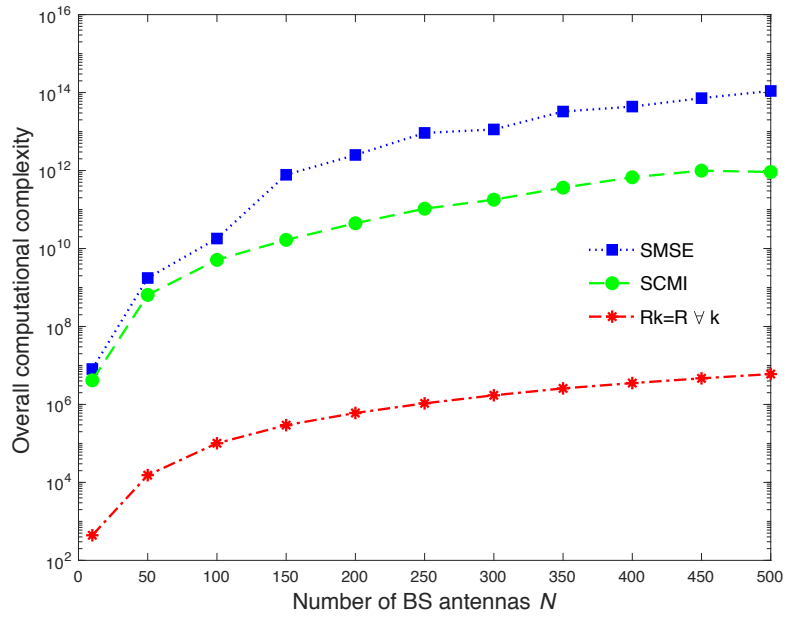


Figure 4.9 Overall computational complexity versus the number of BS antennas N comparing different training sequence methods corresponding to the optimal pilot length in RZF precoding in the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$ and $K = 6$.

4.5 Summary

As explained in Chapter 2 , accurate CSI estimation is essential for massive MIMO systems to design efficient DL precoding. Chapter 2 also demonstrated that in some realistic scenarios in which the channel exhibits a shorter coherence time, the CSI estimation should be determined more frequently. Although gradient-based iterative algorithms were used to optimise the DL training sequences, and hence, meet the requirement for efficient precoding design, the practical implementation of these iterative algorithms is in general limited by their unpredictable convergence behaviour. As such, iterative approaches for the timely estimate of CSI in massive MIMO systems with limited coherence time may be undesired. Additionally, no closed-form solution for the optimum training sequence can be derived based on such iterative techniques.

Therefore, this chapter was focused on developing a non-iterative solution for the DL training sequence design as well as creating a tractable analysis for the CSI estimation problem in an FDD massive MIMO system. This included providing a closed-form solution for the optimum training structure $\mathbf{S}_{\mathrm{p}}$ and the optimum pilot length T_{p}^{*} based on the objective of maximising the achievable sum rate performance while keeping the coherence time fixed. The described essential objectives are achieved in this chapter by considering a special scenario with a common correlation for all users, which makes the analysis tractable. Specifically, this special scenario allows a feasible framework analysis for the problem of DL CSI estimation to be established, and thus, a straightforward design methodology to be achieved. To this end, a low-complexity non-iterative training sequence solution was developed using the eigenvectors of the transmit spatial correlation matrix that correspond to the largest eigenvalues, which is unique for all users. The obtained unique solution aligns the DL sequence into the relevant users' channel directions that contain the significant portion of energy. Furthermore, the optimal DL training sequence length for BF precoding is obtained analytically in a closed-form. The computational complexity of the training sequence design is analysed and compared with the state-of-the-art sequence designs.

Random matrix theory methods along with the P-DoF channel model are used in this chapter to obtain a closed-form analytical expression for the effective SINR, and hence the sum

rate, of BF and RZF precoding with perfect and imperfect channel estimation. Results characterising the system performance for BF and RZF precoding in perfect and imperfect channel estimation and in correlated and uncorrelated Rayleigh fading channels, are also provided. The results demonstrated that the achievable sum rate is not a monotonically increasing function with increasing numbers of transmit antennas N in FDD massive MIMO systems, due to the imperfect DL channel estimation. Furthermore, the results showed that even when the channels are uncorrelated, a feasible sum rate performance with a DL channel estimation can be realised in an FDD massive MIMO system based when the pre-log fraction is considered. Furthermore, excellent agreement between the numerical, analytical and simulated results are obtained. The results showed that the non-iterative training design based on a common correlation archives a significant reduction in computational complexity compared with the state-of-the-art iterative algorithms.

Importantly, the optimisation framework and the theoretical development for the DL training sequence and CSI estimation, which are established in this chapter, will be taken forward and applied to a more general scenario where users exhibit distinct spatial correlations in the next chapter. In particular, the training sequence solution obtained in this present chapter is a special case of the pilot design that will be developed in the next chapter. The proposed non-iterative training sequences can lead to fast CSI estimation, which can be more feasible for FDD massive MIMO communication systems with limited coherence time.

Chapter 5

Superposition Training Sequence Design for FDD Massive MIMO

5.1 Introduction

As discussed in Chapter 2, designing a DL training scheme in an FDD multiuser massive MIMO system with single-stage precoding, and when users exhibit different correlation patterns, is challenging, which to date has necessitated gradient-descent based on iterative algorithms. Although the DL training sequences were optimised in Chapter 3 for the general scenario with multiple heterogeneous user channels, no closed-form solution for finding the optimum training sequences is possible. Additionally, in real-time systems, it is essential to consistently determine and track the channel variations for all active users with an acceptable latency from channel estimation to precoding. As such, the iterative optimisation approaches for DL training sequence design may be infeasible in some channel conditions. A computationally feasible solution for the DL training sequence in an FDD massive MIMO system was developed in Chapter 4 and a tractable framework analysis for the optimum DL training structure was provided, in the special case where users exhibit the same spatial channel correlation. In particular, Chapter 4 showed explicitly how to assign an optimum training sequence to each of the mobile users individually, whereby the effective directions (eigenvectors) based on the transmit covariance matrix of any user is the same for all mobile users. Furthermore, Chapter 4 provided

an analytical closed-form solution for the optimal training sequence length that maximises the achievable sum rate with limited coherence time. However, designing a feasible DL training scheme with a reduced design complexity in the general scenario of multiuser massive MIMO systems with different correlation patterns remains a challenge due to the lack of a closed-form solution. Therefore, the objective of this chapter is to develop an alternative approach to iterative design for the DL training sequence, in realistic propagation conditions where users exhibit district spatial correlations, and which is feasible with limited coherence time.

As explained in Chapter 2, a downlink training sequence design for FDD massive MIMO systems was developed in [51, 52] based on a two-stage-precoding technique, termed as JSDM. The works in [51, 52] exploit correlations in the spatial domain, where the users within each group exhibit the same spatial correlation. In particular, the research in [51, 52] uses a linear superposition of each group correlation matrix to perform the first of two stages of precoding, thus forming a beam for each group. As such, the number of training sequence length in the DL can be scaled linearly with the number of user groups, which can be less than N, resulting in feasible pilot overhead requirement for FDD operation. While the two-stage precoding technique helps to constrain the training sequence length, opportunistic scheduling and sophisticated clustering algorithms of the user groups, and of the users inside each group, are essential, constraining the approach. Furthermore, the research in [51, 52] does not address the challenge of single-stage precoding training sequence design with K independent user covariance matrices, and thus, cannot predict the optimum training sequence length that maximises the sum rate performance with limited coherence time in this preferred scenario.

Building on the framework analysis and theoretical development provided in Chapter 4, this chapter achieves the described objective using a linear superposition principle, in which the DL training sequences are constructed from the eigenvectors of the K independent channel correlation matrices. This is, an advance over Chapter 4 where a common eigenvector columns matrix held for all the users. Based on the superposition training structure, an analytical closed-form solution for the optimum training sequence length that maximises the DL achievable sum rate in the general scenario is provided for the BF precoder.

In this chapter, a P-DoF channel model and a large system limit based on random matrix theory methods are used to provide an analytical approximation of the SINR for BF and RZF pre-

coding in realistic propagation conditions with heterogeneous user channels, and hence, allows the achievable sum rate to be straightforwardly evaluated. Results characterising the achievable sum rate and optimal pilot length based on superposition training design are presented. The proposed superposition training design is evaluated analytically, numerically and by simulation. The impact of increasing the number of BS antennas while keeping the coherence time fixed is examined based on the P-DoF and OR channel models with both the BF and RZF precoders. Comparison between the sum rate of perfect and imperfect DL channel estimation is provided, which allows the effect of DL channel estimation on the achievable sum rate of an FDD multiuser massive MIMO system to be characterised. Comparisons between the achievable sum rates of correlated multiuser channels in which users exhibit independent correlation matrices, and those of uncorrelated Rayleigh fading channels where users have identical covariance matrices are also provided. This chapter also provides comparisons between the achievable sum rates of the UL channel estimation as conventionally used in a TDD system and the DL channel estimation in an FDD system. In addition, comparisons between the achievable sum rates of the low-complexity superposition training design and the state-of-the-art sequences designed based on SCMI [53] and SMSE [56] gradient-based iterative algorithms, which have already been presented in Chapter 3, are conducted. Furthermore, the computational complexity of the superposition sequence design is analysed and compared with the state-of-the-art iterative algorithms.

The numerical results indicate that the sum rate is maximised by using the superposition training sequence design and different spatial channel correlations for the users. Results show that in the P-DoF channel model, the superposition training design achieves almost the same rate performance in comparison with the state-of-the-art SCMI/SMSE sequence designs. It is also demonstrated that in the OR scattering model, the low-complexity training sequence design incurs a small loss in the rate performance compared with the rates achieved by the state-of-the-art SCMI/SMSE sequence designs. Furthermore, the mathematical analysis developed in this chapter using a random matrix theory tightly agrees with the simulated results. Numerical results show that the optimum training sequence length that is analytically optimised for the BF precoder with the P-DoF channel model can also be used to predict the achievable sum rate performance of the RZF precoder. Importantly, the numerical results illustrate that the

analytical solution for the optimal training sequence length with the P-DoF channel model can be effectively used with high accuracy to predict the achievable sum rate performance in the OR channel model without resorting to computationally demanding Monte Carlo simulations, as carried out in Chapter 3. The results additionally demonstrate that the computational complexity of the superposition sequence design is higher than the sequence designed in Chapter 4 when the spatial correlation is identical for all the users, but it is significantly lower than the iterative-based approaches for the sequence design in Chapter 3.

5.2 Sum rate analysis with perfect channel estimation

In order to provide a baseline for comparison with an imperfect channel estimation, this section develops an asymptotically tight approximation of the SINR, and hence the sum rate, for the BF and RZF precoders with perfect channel estimation based on random matrix theory methods. In the first part of this section, an analytical asymptotic approximation of BF and RZF precoding is given for a general channel correlation model based on random matrix theory methods [30]. In the second part of this section, an analytical expression of the SINR that characterises the performance of the BF and RZF precoders in closed-form is developed based on the P-DoF correlation model. This asymptotically tight approximation is developed based upon the framework analysis in Chapter 4 but with an extension to a more general scenario of heterogeneous user channels.

This chapter considers the transmission system model of an FDD massive MIMO with the DL channel, which is presented in Chapter 3 section 3.2. To this end, the DL achievable sum rate with perfect channel estimation C^{PF}, where PF denotes a perfect channel estimation, for an FDD massive MIMO system under consideration can be expressed as

$$\mathrm{C}^{\mathrm{PF}} = \sum_{k=1}^{K} \log_2\big(1 + \mathrm{SINR}_k\big), \tag{5.1}$$

where SINR_k is the associated SINR term at the k-th UT, as given in (3.5), when the BS employs precoding with perfect channel estimation. The random matrix theory in [30] is used to replace the SINR_k term in (3.5) for the BF and RZF precoders, with the asymptotic approximations

provided in this section. Accordingly, the associated DL achievable sum rate $\bar{\mathrm{C}}^{\mathrm{PF}}$ with perfect channel estimation, for an FDD massive MIMO system is given by

$$\bar{\mathrm{C}}^{\mathrm{PF}} = \sum_{k=1}^{K} \log_2\big(1 + \overline{\mathrm{SINR}}_k\big). \tag{5.2}$$

To obtain the analytical results related to the achievable sum rate performance of the BF with general channel model, let $\mathrm{SINR}_k^{\mathrm{BF,PF}}$ denote the SINR_k for BF precoding, as given in (3.5), with perfect channel estimation, where the subscript PF in this context refers to the perfect channel estimation. An asymptotically tight approximation of $\mathrm{SINR}_k^{\mathrm{BF,PF}}$ with perfect channel estimation and based on the scenario where users exhibit different spatial channel correlations reads

$$\overline{\mathrm{SINR}}_k^{\mathrm{BF,PF}} = \frac{\hat{\xi}\big(\mathrm{tr}(\mathbf{R}_k)\big)^2}{\frac{1}{\rho_{\mathrm{d}}} + \hat{\xi}\sum_{i=1}^{K}\mathrm{tr}\big(\mathbf{R}_k\mathbf{R}_i\big)}, \tag{5.3}$$

where $\hat{\xi} = \left(\frac{1}{K}\sum_{k=1}^{K}\mathrm{tr}(\mathbf{R}_k)\right)^{-1}$ corresponds to the precoding normalisation and $\mathbf{R}_k$ is the k-th user's correlation matrix at the BS-side. The SINR expression in (5.3) is generally valid for any type of correlation model. For the P-DoF channel model, the performance of the BF precoder can be characterised in closed-form, which is developed in subsection 5.2.1.

To provide the key results for the RZF precoder with perfect channel estimation under the general channel correlation model, let $\mathrm{SINR}_k^{\mathrm{RZF,PF}}$ denote the SINR_k for the RZF precoding as given in (3.5) with perfect channel estimation, where PF indicates a perfect channel estimation. An asymptotically tight approximation of $\mathrm{SINR}_k^{\mathrm{RZF,PF}}$ for the scenario where users admit distinct spatial correlations is given as

$$\overline{\mathrm{SINR}}_k^{\mathrm{RZF,PF}} = \frac{N\,\hat{\xi}\,\hat{\delta}_k^2}{\frac{(1+\hat{\delta}_k)^2}{\rho_{\mathrm{d}}K} + \hat{\xi}\sum_{i=1}^{K}\left(\frac{1+\hat{\delta}_k}{1+\hat{\delta}_i}\right)^2\hat{\mu}_{k,i}}, \tag{5.4}$$

where $\hat{\delta}_k \in \mathbb{R}$ is an auxiliary variable that can be numerically solved using a fixed-point algorithm. The term $\hat{\xi}$ for the RZF precoding is defined later in (5.12). A closed-form solution to $\hat{\delta}_k$ with perfect channel estimation is provided in Theorem 5.2 of this chapter based on the P-DoF channel model. Using a standard fixed-point algorithm and defining a recursion on integer t for

$t = 1, 2, \ldots,$

$$\hat{\delta}_k^{(t)} = \frac{1}{N}\mathrm{tr}\left(\mathbf{R}_k\left(\frac{1}{N}\sum_{i=1}^{K}\frac{\mathbf{R}_i}{1+\hat{\delta}_i^{(t-1)}} + \zeta\mathbf{I}_N\right)^{-1}\right), \tag{5.5}$$

with an initial value $\hat{\delta}_k^{(0)} = 1/\zeta$ for all k [30], where $\zeta > 0$ is the regularisation parameter, a unique non-negative solution to the variable $\hat{\delta}_k$ is obtained numerically as the limit

$$\hat{\delta}_k = \lim_{t\to\infty} \hat{\delta}_k^{(t)}, \tag{5.6}$$

which is substituted into

$$\hat{\mathbf{T}} = \left(\zeta\mathbf{I}_N + \frac{1}{N}\sum_{k=1}^{K}\frac{\mathbf{R}_k}{1+\hat{\delta}_k}\right)^{-1}, \tag{5.7}$$

to obtain auxiliary random matrix $\hat{\mathbf{T}} \in \mathbb{C}^{N\times N}$ with perfect channel estimation. Auxiliary matrix $\hat{\bar{\mathbf{T}}} \in \mathbb{C}^{N\times N}$ is given by

$$\hat{\bar{\mathbf{T}}} = \hat{\mathbf{T}}\left(\mathbf{I}_N + \frac{1}{N}\sum_{k=1}^{K}\frac{\mathbf{R}_k\hat{\bar{\delta}}_k}{(1+\hat{\delta}_k)^2}\right)\hat{\mathbf{T}}, \tag{5.8}$$

and $\hat{\bar{\boldsymbol{\delta}}} \triangleq [\hat{\bar{\delta}}_1 \ldots \hat{\bar{\delta}}_K]^{\mathrm{T}}$ is given as

$$\hat{\bar{\boldsymbol{\delta}}} = \left(\mathbf{I}_K - \hat{\mathbf{J}}\right)^{-1}\hat{\bar{\mathbf{v}}}, \tag{5.9}$$

where $\hat{\mathbf{J}} \in \mathbb{C}^{K\times K}$ and $\hat{\bar{\mathbf{v}}} \in \mathbb{C}^{K}$ are obtained from the expressions given in (5.10) and (5.11).

$$[\hat{\mathbf{J}}]_{k,l} = \frac{\mathrm{tr}\left(\mathbf{R}_k\hat{\mathbf{T}}\mathbf{R}_l\hat{\mathbf{T}}\right)}{\left(N(1+\hat{\delta}_k)\right)^2}, \qquad 1 \le k, l \le K, \tag{5.10}$$

$$[\hat{\bar{\mathbf{v}}}]_k = \frac{1}{N}\mathrm{tr}\left(\hat{\mathbf{T}}\mathbf{R}_k\hat{\mathbf{T}}\right), \qquad 1 \le k \le K. \tag{5.11}$$

Parameter $\hat{\xi} \in \mathbb{R}$ in (5.4) is determined by substituting $\hat{\mathbf{T}}$ and $\hat{\bar{\mathbf{T}}}$ into

$$\hat{\xi} = \left(\mathrm{tr}(\hat{\mathbf{T}}) - \zeta\mathrm{tr}(\hat{\bar{\mathbf{T}}})\right)^{-1}. \tag{5.12}$$

The auxiliary variable $\hat{\mu}_{k,i} \in \mathbb{R}$ in (5.4) is obtained based on the dominated convergence theorem and the continuous mapping theorem [30]. Consequently, $\hat{\mu}_{k,i}$ is provided in expressions (5.13), (5.14) and (5.15).

$$\hat{\mu}_{k,i} = \frac{1}{N}\mathrm{tr}\big(\mathbf{R}_k\hat{\mathbf{T}}_i'\big) - \frac{2\mathrm{Re}\Big(\mathrm{tr}\Big(\mathbf{R}_k\hat{\mathbf{T}}\Big)\mathrm{tr}\Big(\mathbf{R}_k\hat{\mathbf{T}}_i'\Big)\Big)(1+\hat{\delta}_k) - \mathrm{tr}\Big(\mathbf{R}_k\hat{\mathbf{T}}\Big)^2\hat{\delta}_i'}{\big(N(1+\hat{\delta}_k)\big)^2} \tag{5.13}$$

$$\hat{\mathbf{T}}_i' = \hat{\mathbf{T}}\left(\mathbf{R}_i + \frac{1}{N}\sum_{k=1}^{K}\frac{\mathbf{R}_k\hat{\delta}_k'}{(1+\hat{\delta}_k)^2}\right)\hat{\mathbf{T}} \tag{5.14}$$

$$\hat{\boldsymbol{\delta}}' = \big(\mathbf{I}_K - \hat{\mathbf{J}}\big)^{-1}\hat{\mathbf{v}}' \tag{5.15}$$

where $\hat{\mathbf{v}}' \in \mathbb{C}^K$ denotes

$$[\hat{\mathbf{v}}']_k = \frac{1}{N}\mathrm{tr}\big(\hat{\mathbf{T}}\mathbf{R}_k\hat{\mathbf{T}}\mathbf{R}_k\big), \qquad 1 \le k \le K. \tag{5.16}$$

Comparing (5.4) to (4.4) in Chapter 4, which provides the SINR expression for the RZF precoder with common correlation between users, shows that the users now exhibit distinct spatial correlation matrices at the BS. The SINR expression of the RZF precoder in (5.4) is generally valid for any channel correlation model and provided in terms of several auxiliary variables, which arise from the asymptotic random matrix theory methods. Finding a simplified expression for the SINR with RZF precoding is not straightforward due to the fixed-point algorithms and the several auxiliary variables that require to be numerically solved. Nonetheless, a special case where the performance of the RZF precoder with perfect channel estimation and different spatial correlation matrices can be characterised in closed-form is developed based on the P-DoF channel model. The expressions in (5.3) and (5.4) for the BF and RZF precoding are modified versions of Eq. (23) and Eq. (26) in [30], respectively, where a system model with a single-cell FDD multiuser massive MIMO is considered and the BF and RZF precoders are designed for downlink transmission based on perfect channel estimation.

5.2.1 Sum rate analysis for BF and RZF precoding with the P-DoF channel model and perfect channel estimation

In this subsection, the P-DoF channel model, described in Chapter 2 subsection 2.7.1, is used to obtain a simplified sum rate analysis for the BF and RZF precoders with perfect CSI

estimation when user channels are statistically different and independent of each other with non-overlapping AoA supports. The following Theorem provides a simplified analysis for the SINR of BF precoding with perfect CSI estimation in (5.3), and hence the sum rate, based on the P-DoF channel model.

Theorem 5.1. *An asymptotically tight approximation for the SINR of BF precoding with perfect channel estimation and based on the P-DoF channel model with distinct spatial correlation matrices is given as*

$$\overline{\mathrm{SINR}}^{\mathrm{BF,PF}} = \frac{N}{\frac{1}{\rho_{\mathrm{d}}} + \frac{N}{P}}. \tag{5.17}$$

For the high SNR case, so that $\rho_{\mathrm{d}} \to \infty$, an analytical closed-form solution for the achievable sum rate with perfect channel estimation emerges. In particular, in the special case of high SNR, the achievable sum rate $\bar{C}^{\mathrm{BF,PF},\infty}$, can be further simplified to

$$\bar{C}^{\mathrm{BF,PF},\infty} = K \cdot \log_2(1 + P), \tag{5.18}$$

where P represents the number of degrees of freedom offered by the physical channel.

The proof of Theorem 5.1 is similar to the proof of Theorem 4.1, which is presented in Appendix B. The analysis in Theorem 5.1 shows that the achievable sum rate performance depends only on the number of spatial degrees of the freedom offered by the physical channel. Clearly, the sum rate performance is proportional to the number of degrees of freedom of the channels, which denotes the spatial spreading gain that can be achieved by the massive MIMO channels. In (5.18), the diversity of spatial channel correlations eliminates the interference across multiple users in comparison with (4.15) in Chapter 4, thus maximising the DL achievable sum rate. As shown in equation (4.15), the achievable sum rate is reduced by a factor of K due to the condition of common spatial correlation of the users. The following Theorem characterises the performance of RZF precoding with perfect channel estimation in a closed-form using the P-DoF channel model.

Theorem 5.2. *An asymptotically tight approximation for the SINR for RZF precoding based on the P-DoF channel model is given as*

$$\overline{\mathrm{SINR}}^{\mathrm{RZF,\,PF}} = \frac{N\,\tilde{\xi}\,\tilde{\delta}^2}{\frac{(1+\tilde{\delta})^2}{\rho_{\mathrm{d}}} + \tilde{\xi}\,\tilde{\mu}}, \tag{5.19}$$

where the parameter $\hat{\tilde{\delta}}$ *denotes a closed-form solution to the fixed-point equations arises from* (5.5)–(5.6) *when K independent covariance matrices is used and reads*

$$\tilde{\delta} = \frac{P - 1 - P\zeta + \sqrt{P\zeta^2 + (2+2P)\,P\zeta + 1 - 2P + P^2}}{2P\zeta}, \tag{5.20}$$

where ζ *denotes the regularisation parameter for the RZF precoder and the variables* $\tilde{\xi}$ *and* $\tilde{\mu}$ *are simplified forms of* (5.12) *and* (5.13)*, respectively, and given as*

$$\tilde{\xi} = \frac{\tilde{Q}\zeta\left(\tilde{Q}^2 - P\right)}{\left(1+\tilde{\delta}\right)P^2\,\zeta\left(\tilde{Q} - P\right)}, \tag{5.21}$$

$$\tilde{\mu} = \frac{\left(1+\tilde{\delta}\right)^2 N\,P\left(P^2 + \tilde{Q}^2 - 2\tilde{Q}\,P\right)}{\tilde{Q}^2\left(\tilde{Q}^2 - P\right)}, \tag{5.22}$$

where $\tilde{Q} = 1 + P\zeta + \tilde{\delta}P\zeta$ *is used for notational convenience. The achievable sum rate performance of RZF precoding is obtained by substituting the SINR expression in* (5.19) *into* (5.2)*.*

The simplified expression of the RZF precoder with perfect channel estimation in (5.19) is obtained in similar way to the RZF precoder expression (4.16) in Chapter 4, which is explained in Appendix D. Theorem 5.2 indicates that the SINR of RZF precoding depends on the P, N, regularisation parameter ζ, and average BS transmit power during the data-phase ρ_{d}. This implies that the DL sum rate performance of the RZF with perfect channel estimation can be readily obtained when these relevant system parameters are known. The following section investigates the superposition training sequence design for DL channel estimation in an FDD multiuser massive MIMO system.

5.3 Training sequence design based on linear superposition

The asymptotic sum rate analysis of the BF and RZF precoders in the previous section was developed based on perfect channel estimation at the BS. For practical purposes, this section considers the problem for FDD with imperfect channel estimation using DL training sequences, limited coherence time and when the users exhibit different spatial channel correlations. Consequently, the BS designs the precoders based on $\hat{\mathbf{H}}$, as considered in Chapter 3 section 3.3. To this end, the achievable sum rate, C, for the multiuser massive MIMO systems with imperfect DL channel estimation can be expressed as

$$C = \left(1 - \frac{T_{\mathrm{p}}}{T_{\mathrm{c}}}\right) \sum_{k=1}^{K} \log_2\left(1 + \mathrm{SINR}_k\right), \tag{5.23}$$

where SINR_k is the associated SINR term at the k-th UT, as given in (3.5) with imperfect channel estimation. Comparing (5.23) to (5.1), shows that the achievable sum rate now depends on the channel estimation through the associated SINR at the k-th UT and the pre-log fraction $(1 - T_{\mathrm{p}}/T_{\mathrm{c}})$ term, where the latter corresponds to the amount of time devoted for DL channel estimation.

This chapter builds on the framework analysis of the DL channel estimation developed in Chapter 4, whereby the optimum training sequence is obtained individually for each of the users using the effective eigenvectors of the transmit covariance matrix in the special case of a common spatial correlation for all users. This chapter, however, focuses on finding a feasible solution for DL channel estimation based on a general scenario of single-stage precoding and multiuser massive MIMO systems with a reduced design complexity, thereby extending the training solution provided under the spatial condition of Chapter 4 and avoiding the design of training sequences that require computationally demanding iterative algorithms, such as those presented in Chapter 3, when users admit different spatial correlations. Additionally, based on the superposition training design, this chapter aims to provide the optimum pilot length analytically in closed-form, which maximises the achievable sum rate for a given structure of $\mathbf{S}_{\mathrm{p}}$, for the general scenario under consideration. To this end, the structure of a common pilot matrix $\mathbf{S}_{\mathrm{p}}$ is designed by jointly considering all the K independent channel covariance matrices

where the effective eigenvectors of each of the matrices, as developed in Chapter 4, is combined using the principle of linear superposition. In particular, a unique pilot sequence for the training-phase is developed based on the EVD of the channel covariance matrices

$$\mathbf{R}_k = \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^{\mathrm{H}}, \tag{5.24}$$

where $\mathbf{U}_k = [\mathbf{u}_{k,1}, \ldots, \mathbf{u}_{k,N}] \in \mathbb{C}^{N \times N}$ is a unitary matrix of the eigenvectors and $\mathbf{\Lambda}_k \in \mathbb{R}^{N \times N}$ is a diagonal matrix of the eigenvalues of $\mathbf{R}_k$ arranged in descending order $\lambda_{k,1} \geq \lambda_{k,2} \geq \cdots \geq \lambda_{k,N}$. Specifically, the new spatio-temporal training sequence matrix $\mathbf{S}_{\mathrm{p}} \in \mathbb{C}^{N \times T_{\mathrm{p}}}$ is constructed from the superposition of the first T_{p} eigenvectors of $\mathbf{R}_k$, corresponding to the largest eigenvalues, as expressed in (5.25).

$$\mathbf{S}_{\mathrm{p}} = \frac{\sum_{k=1}^{K} [\mathbf{U}_k]_{:,1:T_{\mathrm{p}}}}{\left\| \sum_{k=1}^{K} [\mathbf{U}_k]_{:,1:T_{\mathrm{p}}} \right\|_{\mathrm{F}}} \tag{5.25}$$

The spatio-temporal common pilot matrix in (5.25) is normalised by the Frobenius norm to satisfy the power constraint $\mathrm{tr}\big(\mathbf{S}_{\mathrm{p}}^{\mathrm{H}} \mathbf{S}_{\mathrm{p}}\big) = T_{\mathrm{p}}$. Finding the optimum training sequence duration T_{p}^{*} that maximises the DL achievable sum rate in an FDD multiuser massive MIMO system with different spatial correlations equates to the optimisation problem defined in (5.26).

$$\begin{aligned} \underset{T_{\mathrm{p}}}{\text{maximise}} \quad & \left(1 - \frac{T_{\mathrm{p}}}{T_{\mathrm{c}}}\right) \sum_{k=1}^{K} \log_2\big(1 + \mathrm{SINR}_k\big) \\ \text{subject to} \quad & 1 \leq T_{\mathrm{p}} \leq \min\big(N, T_{\mathrm{c}}\big) \end{aligned} \tag{5.26}$$

Chapter 4 developed preliminary results for solving the optimisation problem in (5.26) but under a scenario where all users exhibit the same spatial correlation at the BS. In particular, comparing (5.26) to the optimising problem in (4.25), shows that the optimisation problem in (5.26) is now provided in a more general form where each user has an independent channel covariance matrix at the BS. Building on the framework analysis of Chapter 4, a computationally tractable solution to finding an optimum training sequence length T_{p}^{*} in this general setting can be obtained by using large dimension random matrix theory methods. This leads to an analytical solution for T_{p}^{*} in the case of the BF precoder with the P-DoF channel model (see Section 5.5 for further details). The following section elaborates the main analytical results related to the achievable

sum rate of the BF and RZF precoders based on random matrix theory methods [30].

5.4 Achievable sum rate analysis in FDD multiuser massive MIMO systems with imperfect channel estimation

This section provides an accurate approximation for the DL sum rate of the BF and RZF precoders for an FDD massive MIMO system with different spatial channel correlations and imperfect channel estimation. The method of random matrix theory is utilised [30], which allows the numerical optimisation of (5.26) to be achieved with low-complexity. The optimisation problem that maximising the DL sum rate over the training-phase duration in the multiuser massive MIMO system can be rewritten as in (5.27).

$$\begin{aligned} \underset{T_{\mathrm{p}}}{\text{maximise}} \quad & \left(1-\frac{T_{\mathrm{p}}}{T_{\mathrm{c}}}\right)\sum_{k=1}^{K}\log_2\left(1+\overline{\mathrm{SINR}}_k\right) \\ \text{subject to} \quad & 1 \leq T_{\mathrm{p}} \leq \min\left(N, T_{\mathrm{c}}\right) \end{aligned} \tag{5.27}$$

In the following, asymptotic analyses results of the SINR of the BF and RZF precoders are provided with a general correlation model, based on random matrix theory methods [30]. These asymptotic analyses results for the BF and RZF precoders are modified versions of Eq. (23) and Eq. (26) in [30], respectively. Particularly, the precoders under consideration in this section are designed based on imperfect channel estimation in the downlink of a single-cell FDD massive MIMO system. As such, the MMSE channel estimation is obtained based on the downlink training sequence design that is given in (5.25). Consequently, to provide the analytical results for the asymptotic SINR analysis, and hence the sum rate, of the BF with general channel model and imperfect channel estimation, let $\mathrm{SINR}_k^{\mathrm{BF}}$ denote the SINR_k for BF precoding as given in (3.5). An asymptotically tight approximation of $\mathrm{SINR}_k^{\mathrm{BF}}$ with imperfect channel estimation and multiuser massive MIMO based on different correlation matrices reads

$$\overline{\mathrm{SINR}}_k^{\mathrm{BF}} = \frac{\bar{\xi}\left(\mathrm{tr}(\mathbf{\Phi}_k)\right)^2}{\frac{1}{\rho_{\mathrm{d}}}+\sum_{i=1}^{K}\bar{\xi}\,\mathrm{tr}(\mathbf{R}_k\mathbf{\Phi}_i)}, \tag{5.28}$$

where $\bar{\xi} = \left(\frac{1}{K}\sum_{k=1}^{K}\mathrm{tr}(\mathbf{\Phi}_k)\right)^{-1}$ denotes the precoder normalisation and $\mathbf{\Phi}_k$ is the k-th user's covariance matrix with MMSE channel estimation, which is provided in (3.10). Comparing (5.28) to (5.3), shows that the SINR now depends on the MMSE channel estimation and the channel covariance matrix at the BS. The SINR approximation for the BF in (5.28) is generally valid for any channel model and training sequence type. The performance of the BF precoder, and hence the sum rate, can be provided in closed-form using the P-DoF channel model, which is developed later in Section 5.5.

To obtain the performance analysis of the RZF precoder with an imperfect channel estimation and general channel model, let $\mathrm{SINR}_k^{\mathrm{RZF}}$ denote the SINR_k for RZF precoding as given in (3.5). To this end, an asymptotically tight approximation of $\mathrm{SINR}_k^{\mathrm{RZF}}$ for a multiuser massive MIMO system with different spatial correlation matrices and imperfect channel estimation is given as

$$\overline{\mathrm{SINR}}_k^{\mathrm{RZF}} = \frac{N\bar{\xi}\,\delta_k^2}{\frac{(1+\delta_k)^2}{\rho_{\mathrm{d}}K} + \bar{\xi}\sum_{i=1}^{K}\left(\frac{1+\delta_k}{1+\delta_i}\right)^2\mu_{k,i}}, \tag{5.29}$$

where the term $\bar{\xi} \in \mathbb{R}$ is obtained later in (5.37) and $\delta_k \in \mathbb{R}$ is a unique solution to a fixed-point equation that arises from the asymptotic random matrix theory analysis and needs to be numerically solved. An analytical closed-form solution for δ_k with imperfect channel estimation and distinct correlation matrices is provided in Theorem 5.4 (see subsection 5.5.2) based on the P-DoF channel model. Recalling the covariance matrix of the MMSE channel estimate $\mathbf{\Phi}_k$ given in (3.10) and defining a recursion on integer t, where $t = 1, 2, \ldots,$

$$\delta_k^{(t)} = \frac{1}{N}\mathrm{tr}\left(\mathbf{\Phi}_k\left(\frac{1}{N}\sum_{i=1}^{K}\frac{\mathbf{\Phi}_i}{1+\delta_i^{(t-1)}} + \zeta\mathbf{I}_N\right)^{-1}\right), \tag{5.30}$$

with an initial value $\delta_k^{(0)} = 1/\zeta$ for all k [30], where $\zeta > 0$ is the regularisation parameter for the RZF precoder, which is considered to be the inverse of the per-user SNR, $1/\rho_{\mathrm{d}}$ [30]. To this end, the variable δ_k is obtained by the standard fixed-point algorithm as the limit

$$\delta_k = \lim_{t\to\infty}\delta_k^{(t)}. \tag{5.31}$$

After the solution of the fixed-point equations in (5.30) and (5.31) is numerically obtained, it is

substituted into

$$\mathbf{T} = \left(\zeta \mathbf{I}_N + \frac{1}{N}\sum_{k=1}^{K} \frac{\mathbf{\Phi}_k}{1+\delta_k}\right)^{-1}, \tag{5.32}$$

to obtain the auxiliary random matrix $\mathbf{T} \in \mathbb{C}^{N\times N}$ with the covariance matrix of the MMSE channel estimate. Auxiliary matrix $\bar{\mathbf{T}} \in \mathbb{C}^{N\times N}$ is given by

$$\bar{\mathbf{T}} = \mathbf{T}\left(\mathbf{I}_N + \frac{1}{N}\sum_{k=1}^{K} \frac{\mathbf{\Phi}_k \bar{\delta}_k}{(1+\delta_k)^2}\right)\mathbf{T}, \tag{5.33}$$

and $\bar{\boldsymbol{\delta}} \triangleq [\bar{\delta}_1 \dots \bar{\delta}_K]^{\mathrm{T}}$ is given as

$$\bar{\boldsymbol{\delta}} = \left(\mathbf{I}_K - \mathbf{J}\right)^{-1}\bar{\mathbf{v}}, \tag{5.34}$$

where $\mathbf{J} \in \mathbb{C}^{K\times K}$ and $\bar{\mathbf{v}} \in \mathbb{C}^{K}$ are obtained from the expressions given in (5.35) and (5.36).

$$[\mathbf{J}]_{k,l} = \frac{\mathrm{tr}\left(\mathbf{\Phi}_k \mathbf{T} \mathbf{\Phi}_l \mathbf{T}\right)}{\left(N(1+\delta_k)\right)^2}, \qquad 1 \le k,l \le K, \tag{5.35}$$

$$[\bar{\mathbf{v}}]_k \;\; = \frac{1}{N}\mathrm{tr}\left(\mathbf{T}\mathbf{\Phi}_k\mathbf{T}\right), \qquad 1 \le k \le K. \tag{5.36}$$

Parameter $\bar{\xi} \in \mathbb{R}$ in (5.29) is calculated by substituting $\mathbf{T}$ and $\bar{\mathbf{T}}$ into

$$\bar{\xi} = \left(\mathrm{tr}(\mathbf{T}) - \zeta\,\mathrm{tr}(\bar{\mathbf{T}})\right)^{-1}. \tag{5.37}$$

The auxiliary variable $\mu_{k,i} \in \mathbb{R}$ in (5.29) is obtained using the dominated convergence and the continuous mapping theorems as developed in [30], which is modified here for the massive MIMO system under consideration, and thus, it is provided from the expressions given in (5.38), (5.39) and (5.40).

$$\mu_{k,i} = \frac{1}{N}\mathrm{tr}\left(\mathbf{R}_k \mathbf{T}_i'\right) - \frac{2\mathrm{Re}\left(\mathrm{tr}\left(\mathbf{\Phi}_k\mathbf{T}\right)\mathrm{tr}\left(\mathbf{\Phi}_k\mathbf{T}_i'\right)\right)(1+\delta_k) - \mathrm{tr}\left(\mathbf{\Phi}_k\mathbf{T}\right)^2\delta_i'}{\left(N(1+\delta_k)\right)^2} \tag{5.38}$$

$$\mathbf{T}_i' = \mathbf{T}\left(\mathbf{\Phi}_i + \frac{1}{N}\sum_{k=1}^{K}\frac{\mathbf{\Phi}_k\delta_k'}{(1+\delta_k)^2}\right)\mathbf{T} \tag{5.39}$$

$$\boldsymbol{\delta}' = \left(\mathbf{I}_K - \mathbf{J}\right)^{-1}\mathbf{v}' \tag{5.40}$$

where $\mathbf{v}' \in \mathbb{C}^K$ denotes

$$[\mathbf{v}']_k = \frac{1}{N}\mathrm{tr}\big(\mathbf{T}\mathbf{\Phi}_k\mathbf{T}\mathbf{\Phi}_k\big), \qquad 1 \leq k \leq K. \tag{5.41}$$

The SINR expression for the RZF precoder in (5.29) can be applied for any correlation model and training sequence type. Comparing (5.29) to (5.4), shows that the SINR term now depends on the user's covariance matrix with MMSE channel estimate. The SINR approximation for the RZF precoder in (5.29) is given in terms of several auxiliary variables that arise from the asymptotic random matrix analysis. A simplified analysis where the performance of the RZF with imperfect channel estimation and distinct spatial correlation matrices can be characterised in closed-form and is presented in subsection 5.5.2 using the P-DoF channel model. Comparing (5.28) and (5.29) to Eq. (4.27) and Eq. (4.30) in Chapter 4, which provide the SINR approximations for the BF and RZF precoders in an FDD massive MIMO with a common spatial correlation, respectively, shows that the asymptotic approximation for the SINR of BF and RZF precoding of all the users is now statistically different due to distinct spatial channel correlations between users.

5.5 Sum rate and training sequence analysis in the P-DoF channel model for FDD massive MIMO with imperfect channel estimation

Though the optimisation problem formulated in (5.27) may seem straightforward, determining the optimum training sequence length that maximises the sum rate given an arbitrary channel covariance matrices is challenging and forms a key contribution of this research. Nonetheless, the framework analysis developed in Chapter 4 shows that; it is possible to obtain a feasible solution with low-complexity for the optimisation problem in (5.27) based on the random matrix theory and P-DoF channel model. The following subsection investigates a closed-form solution for the optimum training sequence length of a BF precoder using the P-DoF channel model.

5.5.1 Training sequence length optimisation and closed-form sum rate analysis for BF precoding in FDD

Building on the theoretical analysis derived in Chapter 4, this subsection shows how the optimum training sequence length that maximises the DL achievable sum rate is obtained analytically in a closed-form for the BF precoder based on the P-DoF channel model. In particular, the tractable framework analysis developed in Chapter 4 is extended in this subsection to provide a novel closed-form solution for the optimum DL training sequence length of BF precoding in the more general scenario where users exhibit distinct spatial correlation patterns. To this end, the EVD of $\mathbf{R}_k$ in (5.24) is used for training sequence design over which the channel covariance matrices of the users are statistically different and independent of each other with non-overlapping AoA supports. For the particular case of the DFT based P-DoF channel model, described in Chapter 2 subsection 2.7.1, the pilot matrix in (5.25) was constructed based on the principle of linear superposition of subsets formed from the columns of the DFT matrix. To this end, the SINR_k expression for BF precoding, defined by (5.28), can be further simplified for the particular system under consideration. Specifically, following the same framework analysis of the BF precoder in Chapter 4, it is straightforward to obtain simplified expressions for $\text{tr}(\mathbf{\Phi}_k)$ and $\text{tr}(\mathbf{R}_k\mathbf{\Phi}_k)$ in (5.28) based on the P-DoF channel model with independent channel covariance matrices as

$$\text{tr}(\mathbf{\Phi}_k) = \sum_{n=1}^{T_\text{p}} \frac{\lambda_{k,n}^2}{\lambda_{k,n} + K/\rho_\text{p}}, \tag{5.42}$$

$$\text{tr}(\mathbf{R}_k\mathbf{\Phi}_k) = \sum_{n=1}^{T_\text{p}} \frac{\lambda_{k,n}^3}{\lambda_{k,n} + K/\rho_\text{p}}, \tag{5.43}$$

where $\lambda_{k,n} \geq 0\ \forall k = 1,\ldots,K$ and $\forall n = 1,\ldots,T_\text{p}$ are the ordered eigenvalues for the channel covariance matrices $\mathbf{R}_k$ namely, $\lambda_{k,1} = \lambda_{k,2} = \cdots = \lambda_{k,P} = N/P$ and $\lambda_{k,P+1} = \cdots = \lambda_{k,N} = 0$ for each user. The proof of expressions (5.42) and (5.43) is similar to the proof of expressions (4.28), (4.29) in Chapter 4, which is provided in Appendix E. Substituting $\mathbf{R}_k$ of the P-DoF channel model into (3.10) yields, after some algebraic manipulations, the channel estimate co-

variance matrix as

$$\mathbf{\Phi}_k = \frac{1}{\left(1+\frac{K}{\rho_\mathrm{p}}\frac{P}{N}\right)\frac{P}{N}}\mathbf{U}_{k,m}\mathbf{U}_{k,m}^\mathrm{H}, \quad m = \begin{cases} T_\mathrm{p} & \text{if } T_\mathrm{p} \leq P; \\ P & \text{otherwise,} \end{cases} \tag{5.44}$$

where $\mathbf{U}_{k,m} \in \mathbb{C}^{N\times m}$ denotes a matrix constructed from m eigenvectors of $\mathbf{R}_k$. Using the superposition pilot design principle and the same framework analysis in Chapter 4 for the P-DoF channel model, the constraint in the optimisation problem defined in (5.27) can be modified as $T_\mathrm{p} \leq \min(P, T_\mathrm{c})$, which limits the training sequence length by the degrees of freedom offered by the physical channel. In the following, a simplified expression for the SINR term with BF precoding is provided, which is obtained by combining (5.25) and (5.42)–(5.43) with (5.28). Straightforward algebra provides an asymptotic tight approximation for the SINR for the BF precoder with multiuser massive MIMIO system and distinct spatial correlation matrices as

$$\overline{\mathrm{SINR}}^{\mathrm{BF}} = \frac{\rho_\mathrm{d}\rho_\mathrm{p}T_\mathrm{p}}{\left(P/N+\rho_\mathrm{p}\right)\left(KP/N+\rho_\mathrm{d}\right)}, \qquad T_\mathrm{p} \leq P. \tag{5.45}$$

The proof of the SINR expression for the BF precoding with imperfect channel estimation in (5.45) is similar to the proof of the SINR expression (4.41) in Chapter 4, which is presented in Appendix F. Comparing (5.45) to (4.41), shows that an interference-free scenario is obtained in SINR expression (5.45) with the non-overlapping AoA supports due to the independent correlation matrices of the users. In contrast, the interference is maximised in the SINR expression (4.41) due to the condition of common spatial correlation for all the users $\mathbf{R}_k = \mathbf{R}$ and $\mathbf{\Phi}_k = \mathbf{\Phi}$ $\forall k = 1,\ldots,K$. Both SINR terms (5.45) and (4.41) are equal when the number of users $K = 1$ or $\rho_\mathrm{d} = P/N$. The necessary condition for the SINR with the distinct spatial correlation matrices in (5.45) to be higher than the SINR with a common spatial correlation in (4.41) is that $\rho_\mathrm{d} > P/N$ and $K > 1$.

Substituting (5.45) into (5.27) provides a fast numerical optimisation method for the BF precoder. Importantly, at high SNR, i.e. when $\rho_\mathrm{d} = \rho_\mathrm{p} = \rho \to \infty$, analytical solutions emerge. In this case, the average achievable sum rate based multiuser massive MIMO with distinct correlation matrices $\bar{C}^{\mathrm{BF},\infty}$, further simplifies to

$$\bar{C}^{\mathrm{BF},\infty} = \left(1 - \frac{T_\mathrm{p}}{T_\mathrm{c}}\right) \cdot K \cdot \log_2\left(1 + T_\mathrm{p}\right), \;\; T_\mathrm{p} \leq P, \tag{5.46}$$

which leads to Theorem 5.3. The proof of the expression in (5.46) is similar to the proof of the achievable sum rate expression (4.42) in Chapter 4, which is given in Appendix G.

Theorem 5.3. *For K users with distinct spatial correlation matrices, P-DoF channel model and channel coherence time T_c, the novel DL training sequence length T_p^* that maximises the average achievable sum rate, in a multiuser massive MIMO system using BF precoding, at high SNR and with uniform power allocation is*

$$T_\mathrm{p}^* = \begin{cases} P & \text{if } P < \tau, \\ \underset{T_\mathrm{p} \in \{\lceil \tau \rceil \lfloor \tau \rfloor\}}{\arg\max} \; \bar{C}^{\mathrm{BF},\infty} & \text{otherwise}, \end{cases} \tag{5.47}$$

where

$$\tau = \frac{T_\mathrm{c} + 1}{W\big((T_\mathrm{c} + 1)\,\mathrm{e}\big)} - 1. \tag{5.48}$$

The proof of Theorem 5.3 is similar to the proof of Theorem 4.3, which is presented in Appendix H. Similar to the observation characterised previously in Chapter 4 based on Theorem 4.3, the optimum DL training sequence length T_p^* equals the degrees of freedom offered by the physical channel when P is less than τ and saturates at, or below, $\lceil \tau \rceil$, when the DoF $P \geq \tau$ for τ defined in (5.48). This implies that the sum rate of BF precoding with DL multiuser FDD channel estimation can be upper bounded using Theorem 5.3. In particular, given K and T_c, the achievable sum rates for BF precoding with statistically independent channel covariance matrices and $\rho_\mathrm{d} = \rho_\mathrm{p}$ are upper bounded by

$$\bar{C}^{\mathrm{BF},\infty} \leq \left(1 - \frac{\tau}{T_\mathrm{c}}\right) \cdot K \cdot \log_2\left(1 + \tau\right). \tag{5.49}$$

Unlike the achievable sum rate analysis of BF precoding in (4.45) under a common spatial correlation for all the users, an interference-free scenario is achieved in the sum rate expression in (5.49) with K independent channel covariance matrices. Nonetheless, the same behaviour as in Chapter 4 can also be observed for the optimum pilot length that maximises the sum rate. For example, when the DoF exceeds, (5.48), i.e. $P > \tau$, the asymptotic sum rate for the scenario

under consideration, as a function of N, saturates at a constant level below the RHS of (5.49). Importantly, the numerical results also show that the optimum pilot sequence length obtained from Theorem 5.3 for the BF precoder based on the P-DoF channel model is correspondingly similar to those based on a practical OR channel model. The observation means near optimal solutions can be readily found without resorting to computationally intensive optimisation techniques. The achievable sum rate analysis for the RZF precoder based on the analytical P-DoF channel model with imperfect channel estimation is presented in the following subsection.

5.5.2 Sum rate analysis for RZF precoding based on the P-DoF channel model in FDD with imperfect channel estimation

This subsection provides an analytical closed-form solution for the achievable sum rate of the RZF precoder based on the P-DoF channel model and general multiuser scenario where users exhibit distinct spatial correlations. This analysis is an extension to the analysis developed in Chapter 4, where each user exhibited the same correlation at the BS.

Theorem 5.4. *An asymptotically tight approximation for the SINR for RZF precoding based on the P-DoF channel model with K independent channel covariance matrices is given as*

$$\overline{\mathrm{SINR}}^{\mathrm{RZF}} = \frac{N \breve{\xi}\, \breve{\delta}^2}{\frac{(1+\breve{\delta})^2}{\rho_{\mathrm{d}}} + \breve{\xi}\, \breve{\mu}}, \qquad T_{\mathrm{p}} \leq P\cdot \tag{5.50}$$

Parameter $\breve{\delta}$ is a closed-form solution to the fixed-point equations (5.30)–(5.31) *when the covariance matrix with MMSE channel estimation $\mathbf{\Phi}_k$ is of the form* (5.44) *and reads*

$$\breve{\delta} = \frac{T_{\mathrm{p}} - 1 - Z + \sqrt{Z^2 + (2 + 2T_{\mathrm{p}})\, Z + 1 - 2T_{\mathrm{p}} + T_{\mathrm{p}}^2}}{2Z}, \tag{5.51}$$

where parameter Z in (5.51) *is given as $Z = P\zeta\left(1 + \frac{P\,K}{N\rho_{\mathrm{p}}}\right)$, while the variables $\breve{\xi}$ and $\breve{\mu}$ are simplified forms of* (5.37) *and* (5.38)*, respectively, and given as*

$$\breve{\xi} = \frac{Q\zeta\,(Q^2 - T_{\mathrm{p}})}{(1 + \breve{\delta})\, T_{\mathrm{p}}\, Z(Q - T_{\mathrm{p}})}, \tag{5.52}$$

$$\breve{\mu} = \frac{\left(1+\breve{\delta}\right)^2 N\, T_{\mathrm{p}}\left(T_{\mathrm{p}}^2 + \frac{Q^2 Z}{P\zeta} - 2Q\, T_{\mathrm{p}}\right)}{Q^2\left(Q^2 - T_{\mathrm{p}}\right)}, \tag{5.53}$$

where parameter $Q = 1 + Z + \breve{\delta}Z$ *is used for notational convenience.*

The simplified expression of the RZF precoder with imperfect channel estimation in (5.50) is obtained in similar way to the RZF precoder expression (4.46) in Chapter 4, which is explained in Appendix I. From Theorem 5.4, a tight numerical approximation for the achievable sum rate of RZF precoding in an FDD multiuser massive MIMO system can be readily obtained using (5.27) when the relevant system parameters are known. Due to the complexity of the SINR expression, obtaining an analytical solution for the optimum training sequence length, as in the case of BF precoding, is still challenging. However, the SINR is now given in closed-form, which makes evaluation of the DL achievable sum rate in an FDD massive MIMO system numerically straightforward. Therefore, the optimisation problem in (5.27) is now feasible even for a brute-force search. Importantly, the numerical experiments show that the BF and RZF precoders have very similar optimal training sequence lengths and, thus, T_{p}^* can be reliably chosen also for the RZF precoder using Theorem 5.3. In the following section, the numerical results of the analysis for both BF and RZF precoders are presented.

5.6 Numerical results and discussion

This section presents several simulation and theoretical results, which characterise the system performance of the BF and RZF precoders for multiuser massive MIMO communication systems in the general scenario where users exhibit different spatial channel correlations. The impact of increasing the number of BS antennas on the achievable sum rate of a massive MIMO system over a limited coherence time is investigated. The analytical P-DoF and the OR channel models are considered in the performance evaluation. Results that characterise the sum rate performance for perfect and imperfect channel estimation are provided. Comparison between the achievable sum rate of the UL channel estimation as used in a conventional TDD massive MIMO system and the DL channel estimation is also provided. Results that characterise the achievable sum rate performance of the proposed superposition sequence design are

presented and compared with the state-of-the-art sequence designs. Furthermore, the computational complexity of the proposed superposition training design is analysed and compared with the state-of-the-art sequences and the sequence designed in Chapter 4 based on a common spatial correlation between users.

5.6.1 Comparing system performance for perfect and imperfect channel estimation with distinct spatial correlation matrices

To demonstrate the impact of the DL training sequence on the massive MIMO system performance, this subsection compares the achievable sum rate performance of perfect and imperfect channel estimation. Fig. 5.1 plots achievable sum rate versus the number of BS antennas N for the BF and RZF precoders, based on the P-DoF channel model with $P/N = 0.1$, $T_c = 100$ symbols, SNR= 10 dB and $K = 10$ users. For the P-DoF channel model, parameter $P/N = 0.1$ implies relatively strong correlation. Curves for the BF and RZF precoders with imperfect channel estimation are plotted for the computational methods as follows: numerical (BF & RZF) based on equations (5.27), (5.45), (5.50) and simulated (BF & RZF) based on equation (5.26) with the associated SINR term in (3.5). Curves for the BF and RZF precoders with perfect channel estimation are plotted as follows: numerical (BF & RZF) based on equations (5.2), (5.18), (5.19) and simulated (BF & RZF) based on equation (5.1) with the associated SINR term in (3.5) with perfect channel estimation. These computational methods provide validation between the theoretical and simulated performances, which show an excellent agreement throughout.

Fig. 5.1 shows that for both types of precoders, a significant sum rate gain is obtained with the perfect channel estimation in comparison with imperfect channel estimation, as expected. Clearly, the gap between sum rate performances increases when the number of antenna elements N increases. For example, with the BF precoder, the sum rate with imperfect channel estimation saturates at ~35 bit/s/Hz whereas the sum rate with perfect channel estimation continues to rise, reaching ~58.5 bit/s/Hz at $N = 500$. The sum rate for the RZF precoder with imperfect channel estimation increases more rapidly than the BF precoder before reaching a larger peak value of ~75 bit/s/Hz at $N = 150$, and then slightly decreases monotonically reaching ~67 bit/s/Hz at $N = 500$. However, the sum rate for the RZF precoder with perfect channel estimation continues

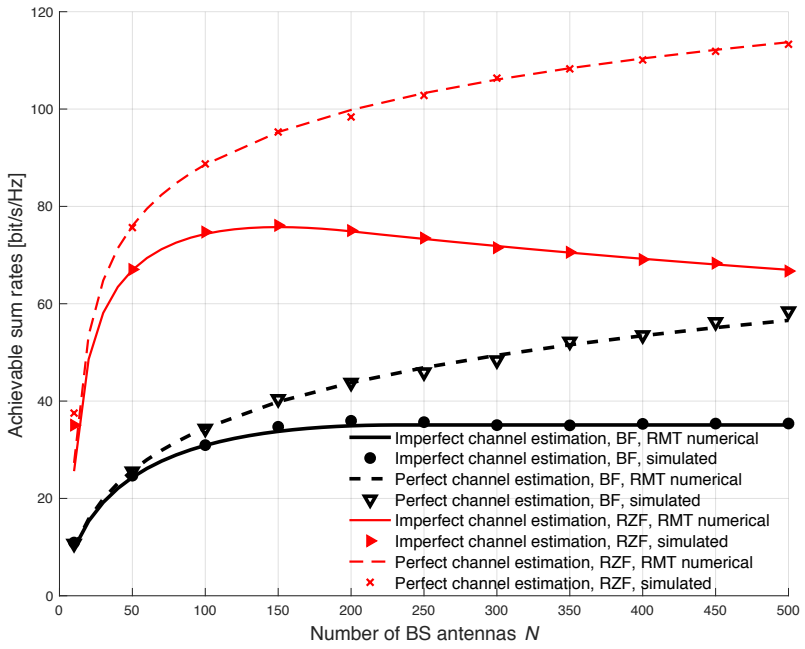


Figure 5.1 Achievable sum rate versus number of BS antennas N based on the P-DoF channel model with $P/N = 0.1$, $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users, comparing perfect and imperfect channel estimation.

to increase with increasing N, reaching a large observed value of ∼113 bit/s/Hz at $N = 500$.

5.6.2 Characterisation of achievable sum rate and training sequence length for distinct spatial correlation matrices

Fig. 5.2 and Fig. 5.3 show plots of the achievable sum rate and the optimum training sequence length T_p^*, respectively, versus the number of BS antennas N, comparing precoder performance in correlated multiuser system and uncorrelated channels[1] where users exhibit a common spatial correlation, when the P-DoF channel model is used with $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users. Curves for the BF and RZF precoders are plotted for three computational methods as follows: numerical (BF & RZF) based on equations (5.27), (5.45), (5.50); analytical BF precoding based on equation (5.46) using the pilot sequence length that is chosen according to Theorem 5.3; and simulated (BF & RZF) based on equation (5.26) with the associated SINR term in (3.5). The curves for the BF and RZF precoders with imperfect channel estimation under uncorrelated channels correspond to those already presented in Chapter 4,

[1]For the P-DoF channel model, parameter $P/N = 1$ corresponds to uncorrelated channels.

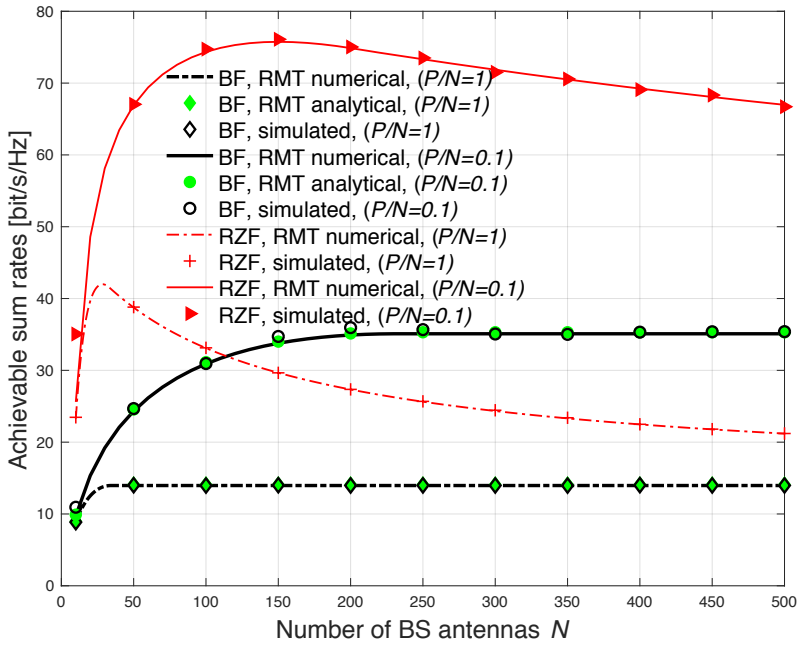


Figure 5.2 Achievable sum rate versus number of BS antennas N for DL channel estimation using the P-DoF channel model with $P/N = \{0.1, 1\}$, $T_c = 100$ symbols, SNR $= 10$ dB and $K = 10$ users.

Fig. 4.2. The reason for carrying out this comparison is to show the effect of the diversity of correlated channels on the achievable sum rate performance in comparison to the scenario of a common correlation for all users.

In Fig. 5.2, the achievable sum rate of BF precoder in the correlated channel with $P/N = 0.1$ increases more gradually than the uncorrelated channel (i.e. $P/N = 1$) before saturating at ~35 bit/s/Hz, though for values of $N > 200$. The saturation of the sum rate for the BF precoder, regardless of the level of correlation in the channel, follows the behaviour predicted by (5.49) for correlated and uncorrelated channels. The sum rate for the RZF precoder in the correlated channel with $P/N = 0.1$ increases more rapidly than the (i.e. $P/N = 1$) uncorrelated channel before reaching a larger peak value of ~75 bit/s/Hz at $N = 150$. For values of $N > 150$, the sum rate slightly decreases monotonically reaching ~67 bit/s/Hz at $N = 500$. As the number of BS antennas N increases, the sum rate decreases due to the residual interference caused by imperfect channel estimation. In particular, as N gets large, the residual interference increases because the capability of the RZF precoder to cancel the interference decreases. The results show that the BF and RZF precoders achieve a significant improvement in the maximum value

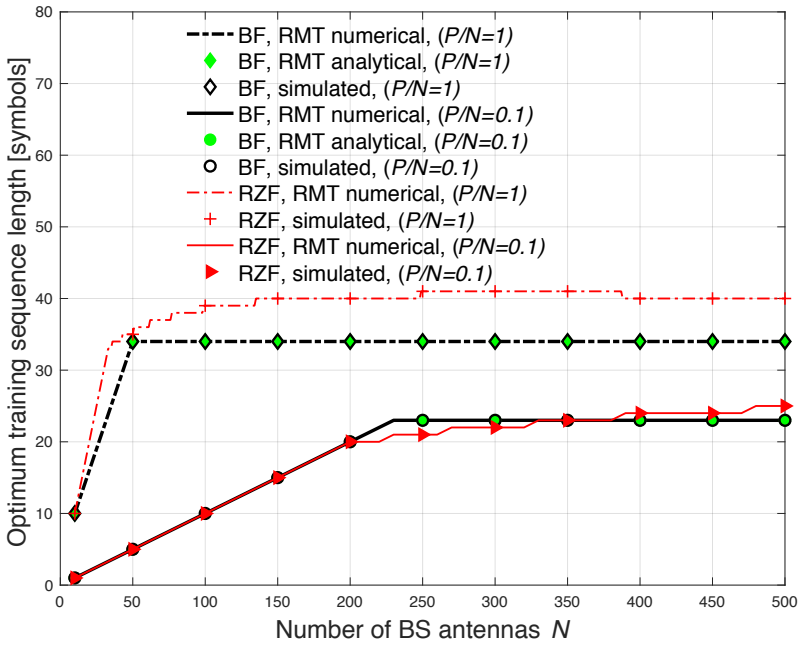


Figure 5.3 Optimum training sequence length T_{p}^{*} versus number of BS antennas N for DL channel estimation using the P-DoF channel model with $P/N = \{0.1, 1\}$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 10$ users.

of achievable sum rate when the channel is strongly correlated.

In Fig. 5.3, for the correlated channel, where $P/N = 0.1$, the optimum training sequence length T_{p}^{*} initially increases linearly with N for both the BF and RZF precoders up to $N = 230$ for the BF precoder and $N = 200$ for the RZF precoder. These regions of the plots corresponds to when $P < \tau$ in (5.47). After these regions, T_{p}^{*} saturates at 23 symbols for the BF precoder while T_{p}^{*} continues to increase slightly for the RZF precoder. The results in Fig. 5.2 & Fig. 5.3 confirm that maximising the sum rate leads to a feasible optimum training sequence length when DL channel estimation is used in an FDD massive MIMO system. The results indicate that the diversity of correlated channels can be effectively exploited in an FDD mutiuser massive MIMO system to maximise the DL achievable sum rate. Furthermore, excellent agreement between the numerical, analytical and simulated results are obtained. Comparing the correlated channels with K independent channel covariance matrices and uncorrelated channels with identical channel covariance matrices $\mathbf{R}_k = \mathbf{I}_{\mathrm{N}}\ \forall k = 1, \ldots, K$, which is provided in Chapter 4, show that the sum rates of the BF and RZF precoders are significantly degraded with the common correlation between users. Nonetheless, a feasible sum rates can be realised at $N < 100$ when

the channel coherence time is limited (i.e. $T_c = 100$ symbols), as shown in Fig. 5.2. This is due to the fact that the sum rates are obtained by optimising the training sequence length through the maximisation of $(1 - T_p/T_c) \sum_{k=1}^{K} \log_2(1 + \text{SINR}_k)$, criterion instead of only minimising the mean square error of the channel estimate criterion, as typically considered in the conventional analyses of FDD and TDD systems.

5.6.3 Comparing system performance for downlink FDD and uplink TDD channel estimation

Having demonstrated the performance of DL training sequence and channel estimation of an FDD multiuser massive MIMO system, it is pertinent to compare the achievable sum rate performance with that obtained for UL channel estimation as conventionally used in a TDD system whereby the uplink-downlink channel with perfect reciprocity is considered. In a TDD based system, a superposition of orthogonal UL training sequences is transmitted by the users to the BS, and the BS estimates the UL channel by using an MMSE channel estimator. For reciprocal channels, the number of pilot symbols required for the UL channel estimation is proportional to the number of users K, which reflects the DoF on the UL.

Fig. 5.4 plots achievable sum rate versus the number of BS antennas N for the BF and RZF precoders, comparing DL superposition training design of an FDD channel estimation and UL channel estimation of a TDD system. The system performances are investigated for the correlated P-DoF channel model with $P/N = 0.1$, $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users. For both types of precoders, a larger sum rate is obtained with an UL training sequence and imperfect channel estimation (i.e., TDD system) than DL training sequence and imperfect channel estimation (i.e., FDD system), as expected. This is due to the fact that the training sequence length saturates in the FDD system with imperfect DL channel estimation as N increases, and thus, the achievable sum rate is decreased. In contrast, in the TDD system, the training sequence length is fixed and independent on N, as explained in Chapter 2, which resulting in increasing the sum rate performance as N increases. The results demonstrate that with the BF precoder the DL sum rate with DL channel estimation of an FDD system saturates at 35 bit/s/Hz whereas the UL sum rate with UL channel estimation of TDD system continues to rise,

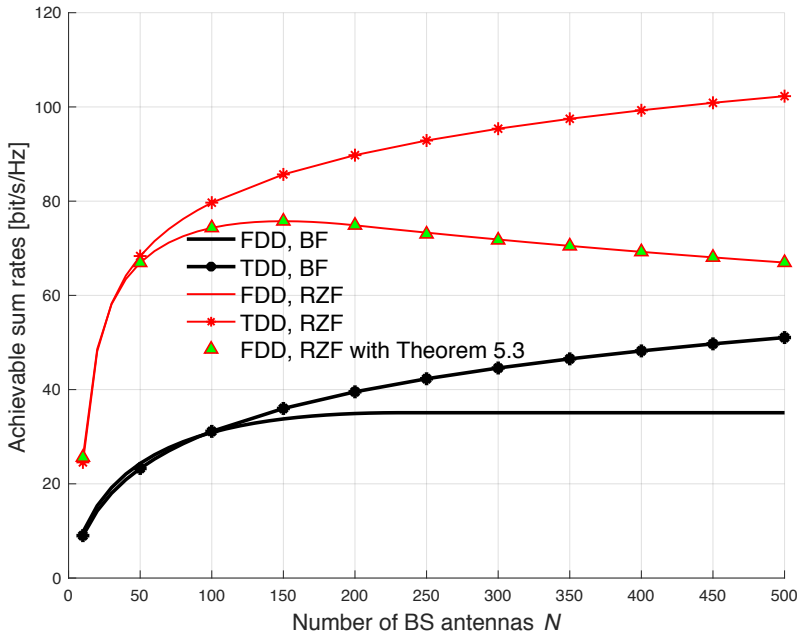


Figure 5.4 Achievable sum rate versus number of BS antennas N based on the P-DoF channel model with $P/N = 0.1$, $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users, comparing DL (i.e. FDD) and UL (i.e. TDD) channel estimation.

albeit slowly, reaching ~51 bit/s/Hz at $N = 500$. The sum rate difference is greater for the RZF precoder with the UL scheme continuing to boost sum rate with increasing N to ~102 bit/s/Hz at $N = 500$ whereas with DL channel estimation the sum rate peaks at ~76 bit/s/Hz for $N = 150$. Nonetheless, for the system parameters considered, over a practical number of antennas from $N = 100$ to $N = 150$, the DL and UL channel estimation schemes offer comparable sum rate performances. The green markers in Fig. 5.4 correspond to the performance of the DL training sequence and channel estimation of an FDD RZF precoder when the training sequence lengths are chosen according to the analytical closed-form solution given in Theorem 5.3 for the BF precoder. Importantly, Fig. 5.4 shows that the closed-form solution for the pilot sequence length mathematically derived in Theorem 5.3 for the DL BF precoder can be accurately used to predict the performance for the DL RZF precoder.

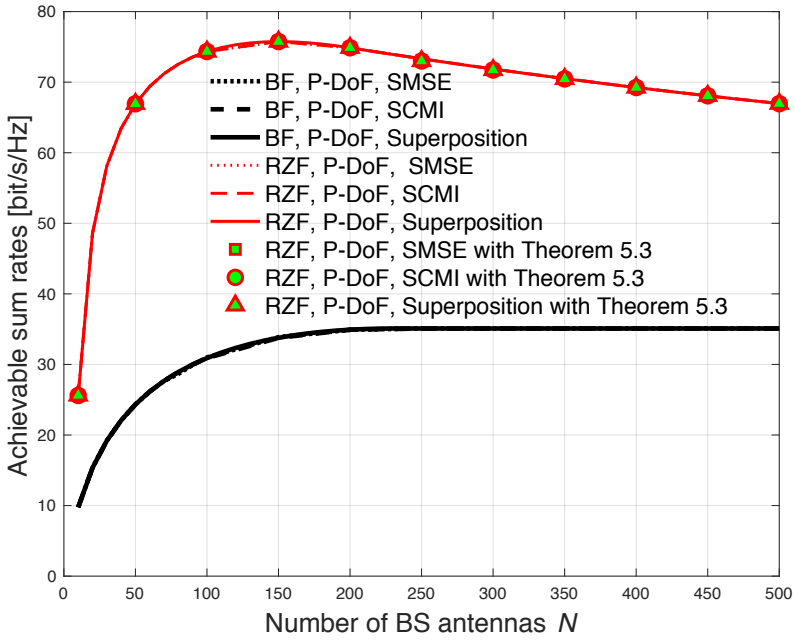


Figure 5.5 Achievable sum rate versus number of BS antennas N comparing different training sequence designs using the P-DoF channel model with $P/N = 0.1$, $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users.

5.6.4 Comparing system performance of the proposed training sequence design and the state-of-the-art sequence designs

Having indicated the feasibility of a DL training sequence based on superposition, this subsection compares the achievable sum rate performance of the superposition sequence design with the SCMI/SMSE sequence designs. The sum rate performance is compared based on both the P-DoF and OR channel models.

5.6.4.1 Performance evaluation of the proposed training sequence design based on the P-DoF channel model

Results in Fig. 5.5 were obtained for the correlated P-DoF channel model with $P/N = 0.1$ whereas the other salient system parameters remain unchanged at $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users. Note that the curves for the superposition sequence design correspond to those already presented in Fig. 5.2. The results in Fig. 5.5 demonstrate that for both types of precoders, all the training sequence designs exhibit essentially the same sum rate performances.

For example, with a BF precoder, the rates saturate at 35 bit/s/Hz, whereas in the RZF precoder the rates peak at ∼76 bit/s/Hz for $N = 150$. The dotted lines for SMSE, and the dash-dotted lines for SCMI are indistinguishable from the solid line of the proposed superposition sequence design for both BF and RZF precoders. In addition, the green markers in Fig. 5.5 represent the performance of the DL RZF precoder when the pilot sequence lengths are chosen according to the closed-form result given in Theorem 5.3 for the DL BF precoder when using DL channel estimation. As can be observed, green markers are indistinguishable from the solid lines (which represent the optimum DL RZF results obtained numerically for all the training sequence designs) demonstrating that the analytical closed-form solution for the optimum pilot sequence length of the BF precoder based on Theorem 5.3 can be reliably selected to predict the achievable sum rate performance of the RZF for all the training sequence designs. This observation also remains valid with the uncorrelated channels, i.e., $P/N = 1$, where the optimum pilot sequence length of the BF precoder in Theorem 4.3 can be used to predict the rate performance of the RZF precoder.

5.6.4.2 Performance evaluation of the proposed training sequence design based on the OR channel model

In order to validate the applicability of the P-DoF channel model, a more practical OR channel model, discussed in subsection 2.7.2, is considered. Though the formulation in (2.9) can accurately predict the rank of the OR channel model that is related to the number of non-zero eigenvalues of $\mathbf{R}_k$, this number may differ between users due to different θ_k. Hence, the maximum number of the effective non-zero eigenvalues across all of the users is selected, i.e. $r = \max_{k=1,\ldots,K}\{r_k\}$, to ensure that all the relevant eigenvectors of each $\mathbf{R}_k$, corresponding to the largest eigenvalues over all the users, are accounted for. When the rank of the OR channel model is obtained, the training sequence length can be reliably selected based on the results derived in Section 5.5.1 for the analytical P-DoF channel model.

Fig. 5.6 and Fig. 5.7 plot the achievable sum rate versus the number of BS antennas N, comparing the proposed superposition training sequence design (5.25) with SCMI/SMSE sequence designs based on the OR scattering channel model for the BF and RZF precoders, respectively.

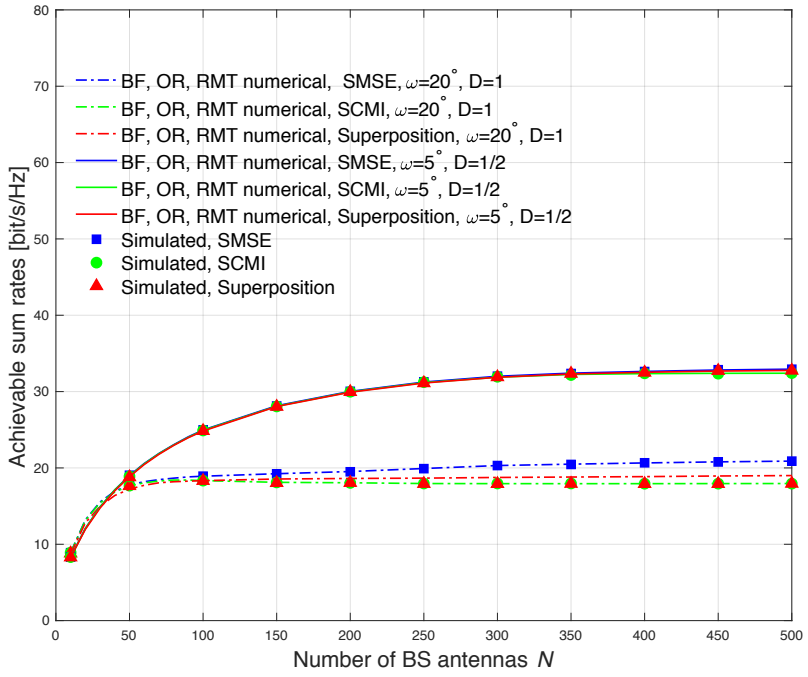


Figure 5.6 Achievable sum rate versus number of BS antennas N, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 10$ users, comparing different training sequence designs based on BF precoder using the OR model with $\omega_{\mathrm{H}} = \{5°, 20°\}$, $D = \{1/2, 1\}$.

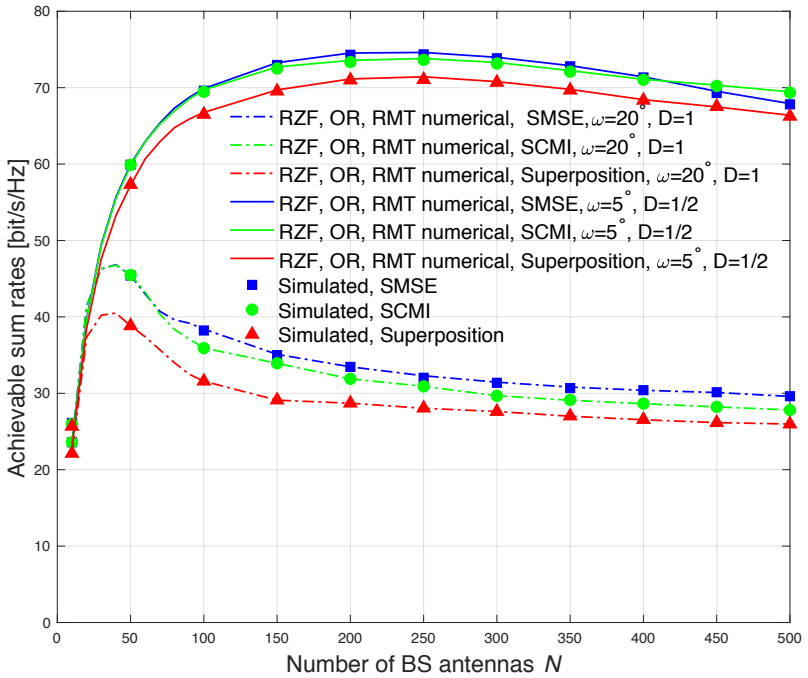


Figure 5.7 Achievable sum rate versus number of BS antennas N, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 10$ users, comparing different training sequence designs based on RZF precoder using the OR model with $\omega_{\mathrm{H}} = \{5°, 20°\}$, $D = \{1/2, 1\}$.

Curves for the BF and RZF precoders are obtained numerically based on equations (5.27), (5.28), (5.29), while the simulated (BF & RZF) precoders are obtained based on equation (5.26) with the associated SINR term in (3.5) and imperfect channel estimation. Fig. 5.6 and Fig. 5.7 are obtained with $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 10$ users. The parameter values for the OR channel model are chosen with angular spread $\omega_{\mathrm{H}} = \{5^{\circ}, 20^{\circ}\}$, normalised antenna spacing $D = \{1/2, 1\}$ and the AoA $\theta_k = \{-57.5^{\circ}, -41.5^{\circ}, -45^{\circ}, -23^{\circ}, -7.5^{\circ}, 7.5^{\circ}, 23.5^{\circ}, 41.5^{\circ}, 45^{\circ}, 57.5^{\circ}\}$ [53, 131]. For the OR model, parameters $\omega_{\mathrm{H}} = 5^{\circ}$, $D = 1/2$ corresponds to strong correlation whereas $\omega_{\mathrm{H}} = 20^{\circ}$, $D = 1$ imply relatively weak correlation. Fig. 5.6 shows almost the same rate performances are obtained in all the training sequence designs for the BF precoder. The solid lines depict numerical analysis based on random matrix theory, while the coloured markers denote simulation. In Fig. 5.7, marginal loss in the rate performance is obtained with the proposed superposition training design in comparison with the state-of-the-art SCMI/SMSE sequence designs. As expected, results in Fig. 5.7 show that the RZF precoder achieves greater sum rate than the BF precoder in Fig. 5.6, both in the correlated and the uncorrelated channels. Also, the results in Fig. 5.6 and Fig. 5.7 demonstrate that significant improvement in the sum rate performances are obtained for both the BF and RZF precoders when the channels are strongly correlated, i.e. $\omega_{\mathrm{H}} = 5^{\circ}$, $D = 1/2$. The results confirm that excellent agreement between numerical and simulated modelling is obtained with the practical OR channel model.

Fig. 5.8 plots achievable sum rate versus N for the BF and RZF precoders showing the impact of using the optimum pilot sequence length obtained by Theorem 5.3 for BF precoder with the P-DoF channel model over the OR scattering channel model. The curves for the BF and RZF precoders in Fig. 5.8 are obtained numerically based on equations (5.27), (5.28), (5.29). The parameter values for the OR model, are $\omega_{\mathrm{H}} = 5^{\circ}$ and $D = 1/2$. The other salient system parameters remain unchanged at $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 10$. The solid lines depict the superposition training sequence design (5.25) and dotted lines and dash-dotted lines correspond to the SCMI/SMSE sequence designs, respectively, while the coloured markers denote the optimum pilot sequence length obtained by Theorem 5.3 for the BF precoder using the P-DoF channel model. The result in Fig. 5.8 confirms that the optimum training sequence length for BF precoder with the P-DoF channel model provides close agreement also in the practical OR model for both the BF and RZF precoders. In particular, it is possible to apply

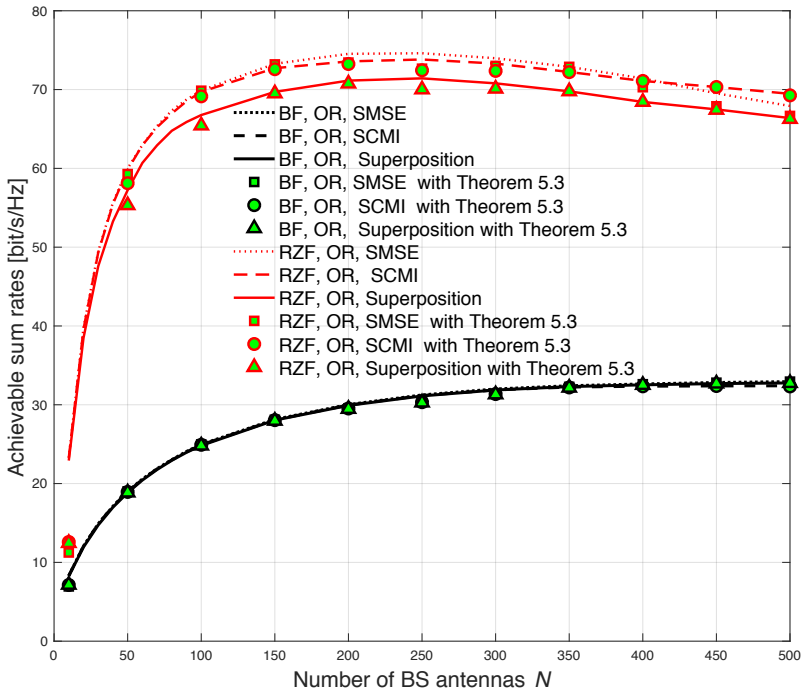


Figure 5.8 Achievable sum rate versus number of BS antennas N, $T_c = 100$ symbols, SNR = 10 dB, and $K = 10$ users, comparing different training sequence designs using the OR model with $\omega_H = 5°$, $D = 1/2$.

Theorem 5.3 to the three training sequence designs in order to obviate the need to search for an optimum training sequence length for all training methods in the more realistic OR channel model. The results illustrate the impact of using Theorem 5.3 for both the BF and RZF precoders over the OR channel, as indicated by the green markers. The results in Fig. 5.8 confirm that the optimum training sequence length for BF precoder with the P-DoF channel model provides close agreement in the practical OR model for both the BF and RZF precoders.

Fig. 5.9 plots achievable sum rate versus N for the BF and RZF precoders demonstrating the impact of applying the optimum pilot sequence length obtained by Theorem 4.3 for BF precoder with the P-DoF channel model over the OR scattering channel model with relatively uncorrelated channels. The curves for the BF and RZF precoders in Fig. 5.9 are obtained numerically based on equations (5.27), (5.28), (5.29). The parameter values for the OR model, are $\omega_H = 20°$ and $D = 1$, which imply relatively weak channel correlations. The other salient system parameters remain unchanged at $T_c = 100$ symbols, SNR = 10 dB and $K = 10$. The result in Fig. 5.9 confirms that the optimum training sequence length for BF precoder with the P-DoF channel model with Theorem 4.3 provides close agreement also in the practical OR model for both the

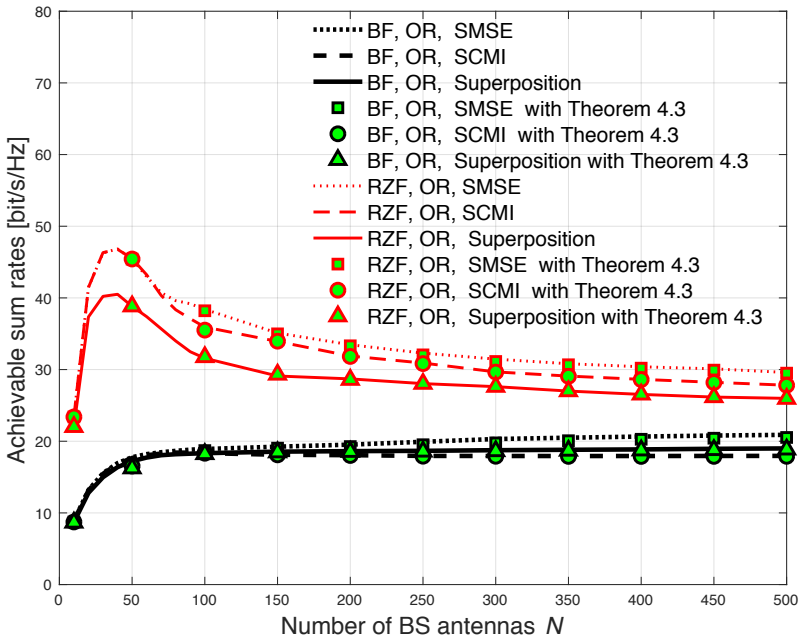


Figure 5.9 Achievable sum rate versus number of BS antennas N, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 10$ users, comparing different training sequence designs using the OR model with $\omega_{\mathrm{H}} = 20°$, $D = 1$.

BF and RZF precoders when the channels are relatively uncorrelated. Overall, Fig. 5.8 and Fig. 5.9 demonstrate the effectiveness of using the P-DoF channel model to provide a practical system design approach, which accurately predicts the performance in more realistic channel models.

Fig. 5.10 and Fig. 5.11 plot the achievable sum rate versus the number of BS antennas N, comparing the performance of the proposed superposition training design in (5.25) for the DL channel estimation of an FDD system with the UL channel estimation of a TDD system in the OR channel model. In particular, Fig. 5.10 and Fig. 5.11 were obtained with $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 10$ users, where the parameter values for the OR channel model are chosen with angular spread $\omega_{\mathrm{H}} = \{2.5°, 5°\}$, and normalised antenna spacing $D = 1/2$. The results in Fig. 5.10 and Fig. 5.11 show that for the system parameters considered, over a practical number of BS array sizes of $N < 200$ antennas, the DL and UL sum rate performances are comparable in the practical OR channel model with correlated channels under both the BF and RZF precoders considered. Specifically, Fig. 5.10 demonstrates that DL channel estimation with the proposed superposition training design can be more beneficial in strongly correlated

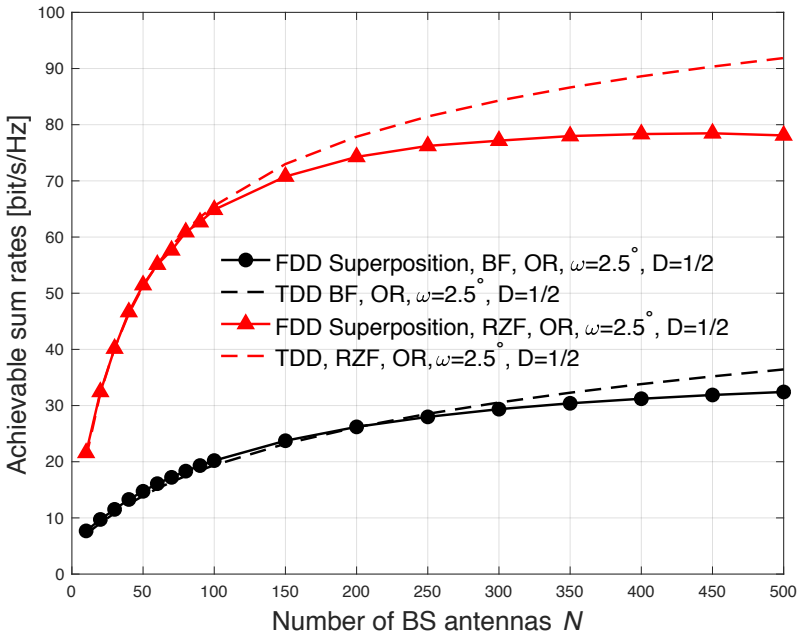


Figure 5.10 Achievable sum rate versus number of BS antennas N based on the OR channel model with $\omega_{\mathrm{H}} = 2.5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR $= 10$ dB and $K = 10$ users, comparing DL (i.e. FDD) and UL (i.e. TDD) channel estimation.

channels.

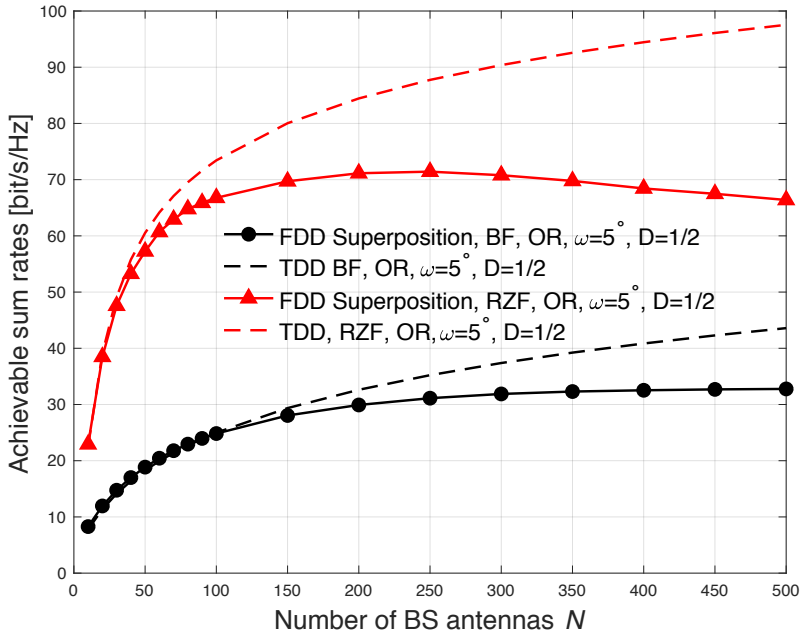


Figure 5.11 Achievable sum rate versus number of BS antennas N based on the OR channel model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR $= 10$ dB and $K = 10$ users, comparing DL (i.e. FDD) and UL (i.e. TDD) channel estimation.

5.6.5 Computational complexity analysis of the superposition training design for FDD

In this subsection, the computational complexity analysis of the superposition training design is presented. The complexity analysis of the superposition design is compared with the state-of-the-art SCMI/SMSE iterative algorithms, which have been already presented in Chapter 3. As explained in Chapter 3, the overall computational complexity is obtained by multiplying the number of iterations each algorithm needs to converge by the number of complex flops involved per iteration. The superposition training design requires only *one iteration*. Table 5.1 summarises the complexity analysis in flops [123] for the superposition training sequence design. The complexity analysis of the SMSE and SCMI sequence designs has been provided in Chapter 3 while the analysis of the training design based on a common spatial correlation (i.e., $\mathbf{R}_k = \mathbf{R}\,\forall$k) has been presented in Chapter 4. Below are the details to explain further how the computational complexity analysis of the proposed superposition sequence design was developed in Table 5.1. Particularly, the analysis is obtained by counting the number of multiplications and additions [123] of each step in the proposed superposition pilot scheme.

- Obtaining the EVD of an $N\times N$ rank deficient matrix in (5.24) for K users needs KN^2r flops.
- In order to obtain the term $\sum_{k=1}^{K}\mathbf{U}_k$ with an $N \times T_{\mathrm{p}}$ matrix in (5.25), $(K-1)(NT_{\mathrm{p}})$ flops are required.
- Obtaining the squared Frobenius norm $||\cdot||_{\mathrm{F}}$ of an $N\times T_{\mathrm{p}}$ matrix in (5.25) requires $2NT_{\mathrm{p}}-1$ flops.
- The scalar matrix multiplication with an $N \times T_{\mathrm{p}}$ matrix needs NT_{p} flops.
- Combining all the flops calculated above, leads to the complexity analysis in Table 5.1.

Fig. 5.12 and Fig. 5.13 show plots of the computational complexity versus the number of BS antennas N, comparing the superposition training design with the SMSE and SCMI training sequence designs, and also with the pilot design based on the common correlation between users. The results in Fig. 5.12 were obtained for the OR model with $\omega_{\mathrm{H}} = 5^\circ$, $D = 1/2$ using

Table 5.1 Computational complexity analysis of the superposition training design

Algorithm	Complexity in flops per iteration
Superposition	$N^2Kr + NT_pK + 2NT_p - 1$

the pilot sequence length for the BF and RZF precoders, respectively. The other salient system parameters remain unchanged at $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users. The results show that the computational complexity of the superposition training design is higher than the sequence designed based on the common correlation between users, as expected. However, the complexity of the superposition sequence design remains significantly lower than the sequence designs based on iterative algorithms. In particular, the results demonstrate that more then a four orders-of-magnitude reduction in computational complexity is achieved using the proposed superposition approach. Hence, signifying the feasibility of the superposition training sequence design for practical implementations compared with state-of-the-art iterative algorithms when users admit distinct spatial correlations. This result is a significant outcome from the thesis.

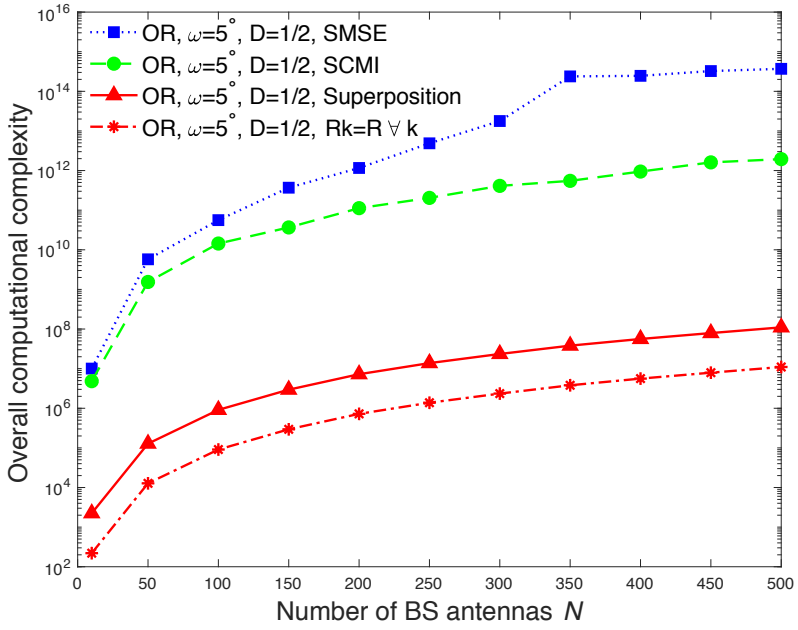


Figure 5.12 Overall computational complexity versus the number of BS antennas N comparing different training sequence methods corresponding to the optimal pilot length in BF precoding in the OR model with $\omega_H = 5°$, $D = 1/2$ and $K = 10$.

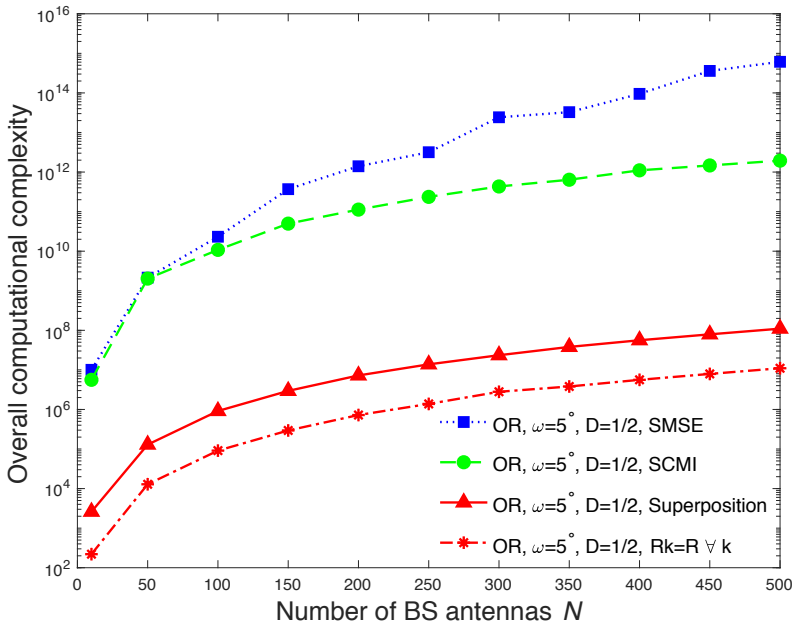


Figure 5.13 Overall computational complexity versus the number of BS antennas N comparing different training sequence methods corresponding to the optimal pilot length in RZF precoding in the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$ and $K = 10$.

5.7 Summary

Building on Chapter 4, which showed that the optimum pilot sequence design that maximises the sum rate is the eigenvectors of the channel covariance matrix, which can be used for all the users since the users exhibit a common spatial correlation, this chapter investigated the training design in the more general scenario where the users have different spatial correlations. To this end, the principle of linear superposition of sequences constructed from the eigenvectors of different users' channel correlation matrices is used to provide a feasible solution for the problem of DL channel estimation in an FDD multiuser massive MIMO system without resorting to computationally intensive iterative algorithms for training designs, such as those considered in Chapter 3. In particular, the essential contribution of this chapter lies in the successful application of the superposition training design to single-stage precoding with K independent channel covariance matrices, with an objective function based on maximising the DL achievable sum rate performance of an FDD multiuser massive MIMO system. Based on the superposition training design, an analytical closed-form solution for the optimum training sequence length T_{p}^{*}

that maximises the DL achievable sum rate $(1-T_{\mathrm{p}}/T_{\mathrm{c}})\sum_{k=1}^{K}\log_2(1+\mathrm{SINR}_k)$ in the BF precoder is provided. Additionally, an analytical asymptotic approximation for the achievable sum rate of the BF and RZF precoders has been provided using large dimension random matrix theory and a P-DoF channel model. The numerical results show that these approximations are accurate for practical, finite values of N and K. Also, the numerical results have been validated with the simulated results given in Chapter 3, which show excellent agreement throughout. Results characterising the system performance for BF and RZF precoders under perfect and imperfect channel estimation when users admit different spatial correlations are presented as well. Furthermore, comparison between the correlated channels with K independent channel covariance matrices and uncorrelated channels with identical channel covariance matrices $\mathbf{R}_k = \mathbf{I}_{\mathrm{N}}$ $\forall k = 1,\ldots,K$ is also provided. While the achievable sum rate performance in the uncorrelated channels (i.e., $\mathbf{R}_k = \mathbf{I}_{\mathrm{N}}$ $\forall k = 1,\ldots,K$) is significantly degraded as the numbers of transmit antennas N increases, nonetheless a feasible sum rate can be realised, e.g., for BS array sizes of $N < 100$. This can be justified by the fact that the optimum training sequence length is obtained by maximising $(1-T_{\mathrm{p}}/T_{\mathrm{c}})\sum_{k=1}^{K}\log_2(1+\mathrm{SINR}_k)$ instead of minimising the mean square error of the channel estimate, as explained in Chapter 4. The results demonstrated that a feasible sum rate can be achieved in a massive MIMO system when finite length pilot sequences are used in the DL estimation of a non-reciprocal wireless channel. Also, an acceptable system performance can be realised for practical BS array sizes of $N < 250$ antennas. Further, the results show that the diversity of spatial channel correlations between multiple users substantially enhances the achievable sum rate performance. In particular, it is demonstrated that for practical BS array sizes of $N < 200$ antennas, the sum rate of a massive MIMO system using DL channel estimation with superposition sequence design is comparable to the performance of a TDD system for both types of precoders considered in a strongly correlated channel. In addition, the numerical results show that the proposed training sequences offer comparable sum rate performances to the state-of-the-art sequences while reducing the computational complexity substantially. Importantly, this chapter demonstrated that the optimum T_{p}^{*} that is analytically optimised for the BF precoder is sufficient to predict the achievable sum rate performance of the RZF precoder, which remains near optimal. This observation also remains valid with the practical OR scattering channel model with only a small loss in performance. These findings

lead to large reductions in the computational complexity of the proposed approach. Overall, the proposed design paradigm opens up the possibility for FDD massive MIMO systems operating in a general scenario of single-stage precoding and more realistic conditions about channel correlation with limited coherence time.

The next chapter investigates low-complexity techniques to further enhance the DL training design in FDD massive MIMO systems based on linear superposition. Furthermore, The next chapter studies if the superimposition pilot framework introduced in this chapter for the ULA also works with the uniform planar array (UPA).

Chapter 6

Techniques to Enhance the Performance of the Training Sequence Design for FDD Massive MIMO

6.1 Introduction

In Chapter 5, a non-iterative DL training sequence design was proposed to provide a feasible solution for the problem of DL channel estimation in the FDD massive MIMO systems with a single-stage precoding unit and limited coherence time. In particular, a low-complexity DL training solution was developed based on the principle of linear superposition, which allows a significant reduction in the complexity compared with the state-of-the-art sequence designs that require computationally demanding iterative algorithms when users exhibit different spatial correlations. Furthermore, the analysis of the results in Chapter 5 indicated that even when the channels are uncorrelated, a feasible sum rate performance with a DL channel estimation can be realised in an FDD massive MIMO system based on the proposed approach when the pre-log fraction is considered. These results motivate the research in developing tractable approaches for improving the system performance based on the non-iterative training sequence design, thus, enhancing the DL achievable sum rate of the FDD massive MIMO systems with limited coherence time. This chapter also studies the use of linear superposition training design,

developed in Chapter 5, in other practical antenna array configurations, such as the UPA.

In the first part of this chapter, the achievable sum rate performance with a DL channel estimation is optimised with respect to different levels of a non-uniform power allocation between the training and data signals. In the second part, a novel weighted superposition for designing a DL training sequence is proposed to enhance the design performance of the conventional superposition, which improves the achievable sum rate of an FDD massive MIMO system when the users exhibit distinct spatial channel correlations. Finally, the chapter investigates the feasibility of applying the linear superposition pilot framework for the DL channel estimation in the UPA configuration at the BS.

The analytical SINR expressions for the BF and RZF precoders, which were developed in Chapter 4 and Chapter 5 based on the asymptotic random matrix theory methods, are adopted in this chapter. Results characterising the achievable sum rate performance of the non-uniform power allocation are presented in this chapter and also compared with the uniform power allocation of the training and data phases in the FDD massive MIMO systems. In addition, the achievable sum rate results of the proposed weighted superposition training sequence design is compared with the rates of the conventional unweighted superposition design. Furthermore, comparisons are provided between the sum rates of the proposed superposition training design for the UPA at the BS and the rates obtained based on the state-of-the-art sequence designs with the SCMI [53] and SMSE [56] iterative algorithms.

The numerical results demonstrate that optimising the energy between the training and data transmissions, enhances the sum rate performance of the RZF precoder in comparison to the uniform power allocation. The analyses of the results in this chapter also show that the weighted superposition training sequence design improves the sum rate performance marginally, especially with relatively weak correlated channels. Importantly, the results additionally indicate that when the BS employs the UPA, the proposed non-iterative superposition training design achieves almost the same rate performance as the state-of-the-art iterative algorithms for DL training sequence designs in the BF precoder, while in the RZF precoder, the superposition design incurs a some loss in the rate in comparison to the rates achieved by the iterative algorithms.

6.2 Energy optimisation between the training and data phases in FDD massive MIMO systems

The results presented in Chapters 3, 4 and 5 are obtained based on the condition that the total available transmit power is divided equally between the training and data signals. In this section, the total available energy between the training-phase and the data-phase is allowed to vary in the DL with an objective to maximise the achievable sum rate of an FDD massive MIMO system under a limited coherence time. This chapter considers a single-cell transmission system model of an FDD massive MIMO with the imperfect DL channel estimation, which is introduced in Chapter 3 section 3.2. The overall coherence time is defined by $T_{\mathrm{c}} \in \mathbb{Z}^{+}$ and enumerated in symbols per transmission block, as described in Chapter 3. Particularly, this section considers a base station that is equipped with a uniform linear array of N antennas, which serves K single antenna UTs. In physics, the power is defined as the amount of energy divided by the time. Let E denote the total transmit energy at the BS, which is obtained by multiplying the overall coherence time T_{c} by the total available transmit power ρ (SNR) as given in (6.1),

$$E = T_{\mathrm{c}}\,\rho, \tag{6.1}$$

where T_{c} is the overall channel coherence time that is divided into training duration T_{p} and data transmission duration T_{d}. In particular, during the training-phase, the BS transmits training sequences of length T_{p} symbols per transmission block for DL channel estimation and the remaining symbols correspond to data transmission, as explained in Chapter 2. At the same time, the total transmit energy available at the BS can also be freely distributed between the pilot and data phases. Accordingly, the energy allocated to the pilot transmission E_{p} for the massive MIMO system under consideration can be expressed as

$$E_{\mathrm{p}} = \varsigma E, \tag{6.2}$$

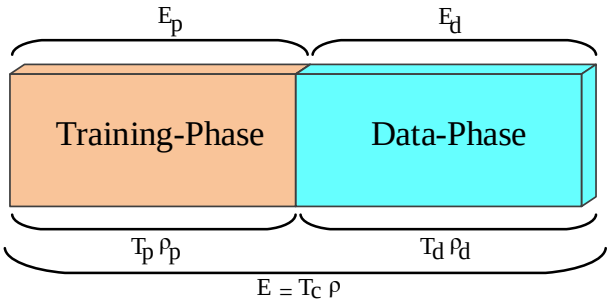


Figure 6.1 Block diagram that characterises the energy distribution between the pilot and data transmissions.

where $\varsigma \in (0, 1]$ denotes the fraction of energy devoted to the pilot transmission. Consequently, the remaining portion of energy that can be allocated for the data transmission E_{d} is given by

$$E_{\mathrm{d}} = (1 - \varsigma)E. \tag{6.3}$$

Consequently, $\rho_{\mathrm{p}} = E_{\mathrm{p}}/T_{\mathrm{p}}$ and $\rho_{\mathrm{d}} = E_{\mathrm{d}}/T_{\mathrm{d}}$, where ρ_{p} and ρ_{d} are the average transmitted power (i.e., SNR relative to unity noise) during the training-phase and the data-phase, respectively. Fig. 6.1 illustrate how the energy can be distributed between the training-phase and the data-phase. The uniform power allocation means equal energy distribution between the pilot and data transmissions. In what follow, the impact of optimising the energy between the training and data transmissions on the achievable sum rate of an FDD massive MIMO system with a limited coherence time is investigated.

6.2.1 The methodology of energy optimisation

Building on the theoretical analysis derived in Chapters 4 and 5, this subsection explains how the energy optimisation results are obtained. In particular, the relation between the energy optimisation and the sum rate maximisation problem of an FDD massive MIMO system with limited coherence time is described.

The pilot energy optimisation is required to improve the CSI estimation accuracy in the DL. This implies that allocating more power to the pilot during the training-phase results in improving the CSI estimation accuracy. The obtained training sequences should satisfy the energy constraint in the sum rate maximisation/optimisation problem, i.e., the pilot matrix is subject to

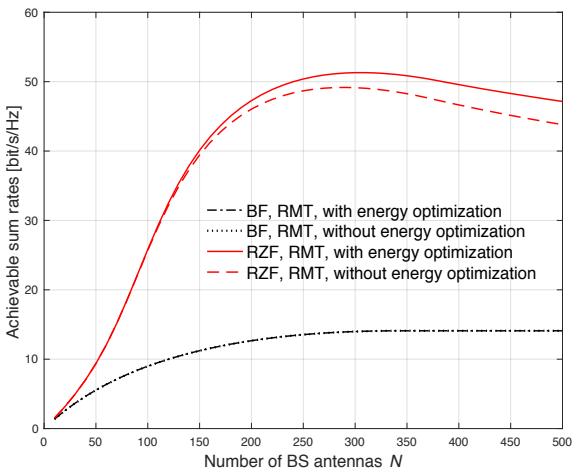
$\mathrm{tr}\left(\mathbf{S}_{\mathrm{p}}^{\mathrm{H}}\mathbf{S}_{\mathrm{p}}\right) = T_{\mathrm{p}}\,\rho_{\mathrm{p}}$. Clearly, $T_{\mathrm{p}}\,\rho_{\mathrm{p}} = E_{\mathrm{p}}$, which corresponds to the amount of energy devoted for DL training-phase. The remaining energy would be allocated to the data transmission.

The simplified analytical expressions of the effective SINR with BF and RZF percodering, which have been developed in Chapter 4 and Chapter 5 using a random matrix theory, are used here. Specifically, the SINR expressions of BF and RZF precoding in (4.41) and (4.46) are used to obtain the results of energy optimisation under a common correlation. In addition, the SINR expressions of BF and RZF precoding in (5.45) and (5.50) are used to obtain the results of energy optimisation with distinct spatial correlations. The simplified analyses based on the random matrix theory allow a tractable numerical energy optimisation with low-complexity.

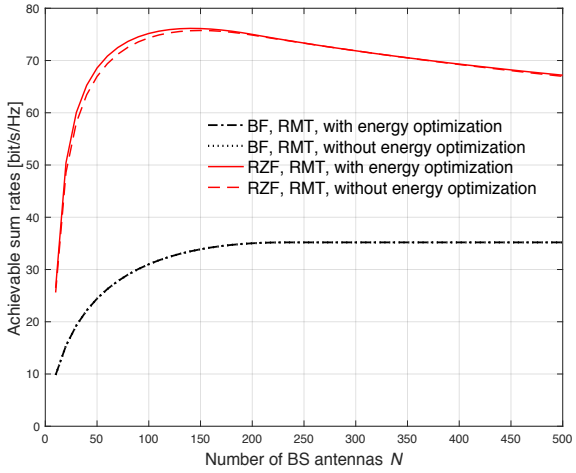
The achievable sum rate optimisation problems in (4.26) and (5.27) are used to obtained the results here. In particular, the optimum power allocation that maximises the achievable sum rate performance with each number of transmit antenna N is obtained by exhaustively searching through all possible values of ς. For each value of N, the rate is optimised with respect to an appropriate weighted value of the energy, which maximises the achievable sum rate performance in an FDD massive MIMO system. It worth noting that optimising the energy between the training and data transmissions would improve the DL achievable sum rate logarithmically. As explained in Chapters 4 and 5, increasing T_{p} allows for more pilot signal energy to be received, but it comes at the cost of reduced achievable sum rate due to a shorter data transmission phase. The following subsection presents the numerical results for the BF and RZF precoders based on the DL channel estimation and energy optimisation in the FDD massive MIMO systems with limited coherence time.

6.2.2 Performance evaluation of the energy optimisation in FDD massive MIMO systems

In this subsection, numerical results that characterise the achievable sum performances of the BF and RZF percoders with the non-uniform and uniform power allocations between the training and data phases, are provided. In the following results, curves for the BF and RZF precoders with energy optimisation refer to a non-uniform power allocation between the pilot and data transmissions whereas curves without energy optimisation denote a uniform power



(a) Common spatial correlation matrix.

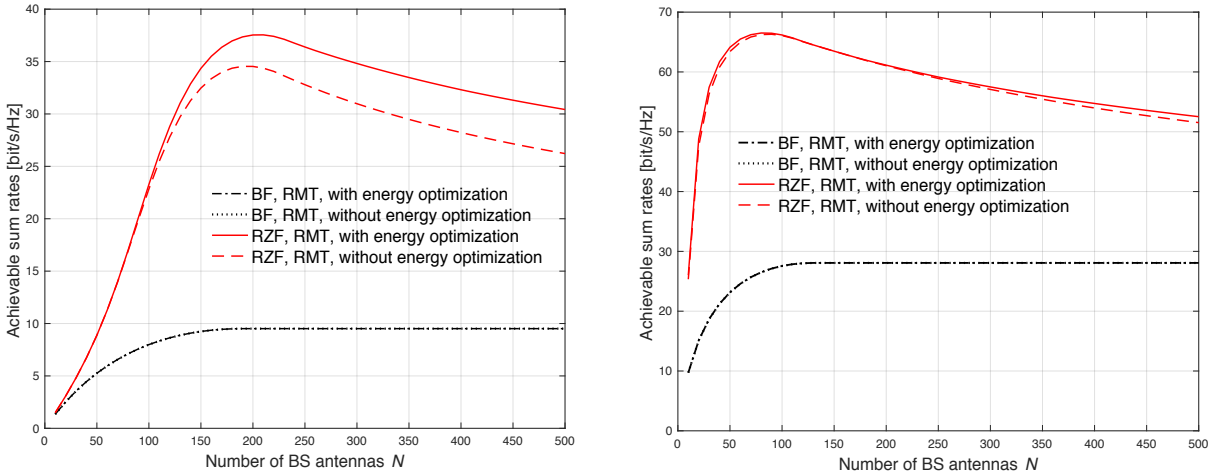


(b) Distinct spatial correlation matrices.

Figure 6.2 Achievable sum rate versus number of BS antennas N, comparing different energy allocation schemes using the P-DoF channel model with $P/N = 0.1$, SNR $= 10$ dB, and $K = 10$ users, and $T_c = 100$ symbols.

allocation. The results are presented for the P-DoF channel model when the salient system parameters P, N, T_c, SNR and K are considered. In addition, the results are provided for the scenarios when the users exhibit common and distinct spatial channel correlations. This subsection uses a uniform linear array in the performance evaluation. Curves for the BF and RZF precoders with common correlation matrix are obtained numerically based on equations (4.26), (4.41) and (4.46), while curves for the BF and RZF precoders with different correlation matrices are obtained numerically based on equations (5.27), (5.45) and (5.50). These analyses for the BF and RZF precoders have already been validated in Chapter 4 and Chapter 5 based on the Monte Carlo simulations, which show an excellent agreement throughout.

Fig. 6.2 plots achievable sum rate versus the number of BS antennas N, comparing the BF and RZF precoders performances based on the P-DoF channel model with $P/N = 0.1$, SNR $= 10$ dB, $K = 10$ users and $T_c = 100$ symbols. For the P-DoF channel model, parameter $P/N = 0.1$ implies relatively strong correlation. As can be observed, when the BS employs BF precoding, almost the same sum rate performances are obtained based on the non-uniform and uniform power allocations. The results in Fig. 6.2a with common spatial correlation show



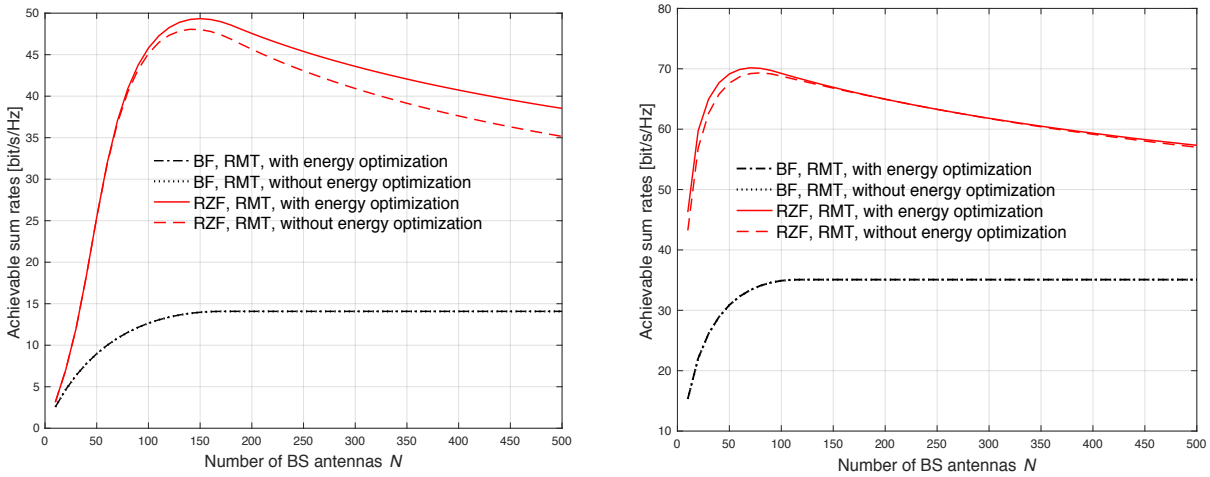
(a) Common spatial correlation matrix.

(b) Distinct spatial correlation matrices.

Figure 6.3 Achievable sum rate versus number of BS antennas N, comparing different energy allocation schemes using the P-DoF channel model with $P/N = 0.1$, SNR $= 10$ dB, $K = 10$ users, and $T_c = 50$ symbols.

that the energy optimisation achieves almost 9 bit/s/Hz improvement in the DL sum rate performance of RZF precoding compared with the uniform power allocation when $N > 150$.

Fig. 6.3 investigates the achievable sum rate performance of the non-uniform power allocation when the channel exhibits a shorter coherence time, i.e., $T_c = 50$ symbols. The other system parameters are $P/N = 0.1$, SNR $= 10$ dB, and $K = 10$ users. The results in Fig. 6.3 show that for the BF precoder, almost the same achievable sum rate performances are obtained based on the non-uniform and uniform power allocations. For the RZF precoder, however, a considerable improvement in the sum rate performance is obtained when the energy between the data and pilot phases is optimised. Comparing the results in Fig. 6.3 to Fig. 6.2, demonstrates that the achievable sum rate performances of the BF and RZF precoders are decreased when a shorter coherence time $T_c = 50$ symbols is considered, as expected. This can be justified by the fact that the available coherence time would be largely occupied by the CSI estimation, leaving insufficient time for transmitting useful data to the users. The results in Fig. 6.3 implies that the energy optimisation when $T_c = 50$ symbols is small, meanwhile the channels are rapidly varying, is worth considering.

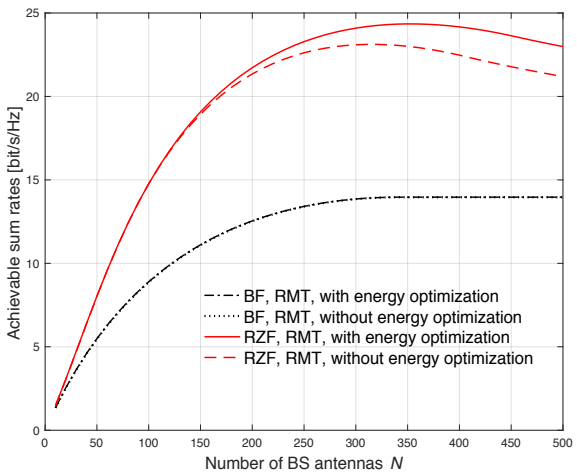


(a) Common spatial correlation matrix.

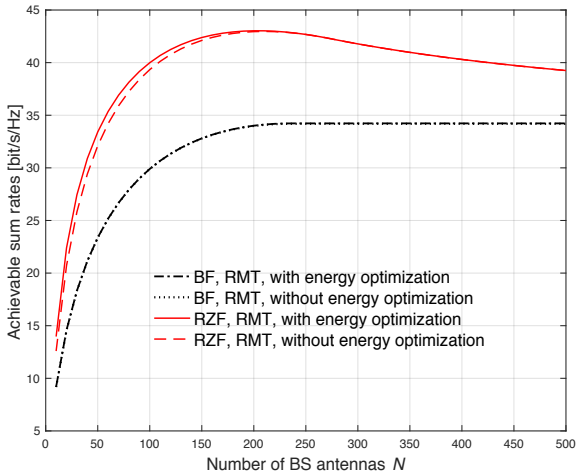
(b) Distinct spatial correlation matrices.

Figure 6.4 Achievable sum rate versus number of BS antennas N, comparing different energy allocation schemes using the P-DoF channel model with $P/N = 0.2$, SNR $= 10$ dB, $K = 10$ users, and $T_c = 100$ symbols.

Fig. 6.4 examines the achievable sum rate performance of the BF and RZF precoders with the non-uniform power allocation when the degrees of correlation in the channel is relatively increased, i.e., $P/N = 0.2$. The other system parameters are SNR $= 10$ dB, $K = 10$ users and $T_c = 100$ symbols. As can be observed in Fig. 6.4a with common spatial correlation, when the level of channel correlation increases, i.e., $P/N = 0.2$, the achievable sum rates of the RZF precoder increase rapidly reaching a peak value of ~50 bit/s/Hz and ~48 bit/s/Hz at $N = 150$ with the non-uniform and uniform training-data power allocations, respectively. For values of $N > 150$, the sum rates slowly decrease monotonically, plateauing at ~39 bit/s/Hz and ~35 bit/s/Hz for $N = 500$, with the non-uniform and uniform power allocations, respectively. For the BF precoder in Fig. 6.4a with common correlation matrix, the achievable sum rates for the non-uniform and uniform power allocations increase steeply with N before saturating at about 14 bit/s/Hz for values of $N > 150$. For the RZF precoder in Fig. 6.4b with distinct spatial correlation matrices, the maximum achievable sum rates are obtained when the number of BS antennas is about $N \sim 70$ for the non-uniform and uniform power allocations. For the BF precoder in Fig. 6.4b with distinct spatial correlation matrices, the achievable sum rates with the



(a) Common spatial correlation matrix.



(b) Distinct spatial correlation matrices.

Figure 6.5 Achievable sum rate versus number of BS antennas N, comparing different energy allocation schemes using the P-DoF channel model with $P/N = 0.1$, SNR $= 0$ dB, and $K = 10$ users, and $T_{\mathrm{c}} = 100$ symbols.

non-uniform and uniform power allocations increase steeply with N before saturating at about 35 bit/s/Hz for values of $N > 100$. The results in Fig. 6.4 demonstrate the effectiveness of optimising the energy between the training and data transmissions on the achievable sum rate performance.

In order to characterise the achievable sum rate performance of the energy optimisation between the training and data transmissions when the total transmit power is decreased, Fig. 6.5 considers a reduced transmit power, i.e., SNR $= 0$ dB. In particular, Fig. 6.5 plots the achievable sum rate versus the number of BS antennas N, comparing the BF and RZF precoders performances based on the P-DoF channel model with $P/N = 0.1$, $T_{\mathrm{c}} = 100$ symbols, SNR $= 0$ dB and $K = 10$ users. The results in Fig. 6.5 show that under both scenarios of common and distinct spatial correlations, a noticeable improvement in the sum rate performances of the RZF precoder are achieved when the energy between training and data transmissions is optimised. Further, the gap between the sum rate performances of BF and RZF precoding decreases when a low SNR value is considered.

6.3 Weighted superposition training design for FDD massive MIMO systems

This section focuses on improving the performance of the low-complexity superposition training design, and hence the DL sum rate of the FDD massive MIMO systems, in realistic propagation conditions with the OR channel model. In particular, the analyses of the results in Chapter 5 showed that in a more realistic OR channel model, the sum rate performance of linear superposition training design incurs a small loss compared with the rates obtained based on the state-of-the-art iterative algorithms for the sequence designs. Here, a novel technique is investigated to improve the performance of the superposition training design. To this end, a weighted superposition training design is proposed with the aim being to enhance the sum rate performance of an FDD massive MIMO system with limited coherence time.

Recalling the eigenvalue decomposition of the channel covariance matrices given in (5.24), with a unitary matrix $\mathbf{U}_k = [\mathbf{u}_{k,1}, \ldots, \mathbf{u}_{k,N}] \in \mathbb{C}^{N\times N}$, which is a complex matrix whose columns are the eigenvectors of the channel covariance matrix $\mathbf{R}_k$, and $\mathbf{\Lambda}_k \in \mathbb{R}^{N\times N}$, which is a diagonal matrix of the eigenvalues of the channel covariance matrix $\mathbf{R}_k$ whose diagonal entries are arranged in descending order $\lambda_{k,1} \geq \lambda_{k,2} \geq \cdots \geq \lambda_{k,N}$. Let $\alpha_{k,t}$ denote the weighted factor, which favours the largest eigenvalues across multiple users as

$$\alpha_{k,t} = \frac{\lambda_{k,t}}{\sum_{k=1}^{K} \lambda_{k,t}} \qquad k = 1, 2, \ldots, K;\ t = 1, 2, \ldots, T_{\mathrm{p}}. \tag{6.4}$$

Let $\mathbf{F}_k \in \mathbb{C}^{N\times T_{\mathrm{p}}}$ be a matrix that contains the first T_{p} weighted eigenvectors of $\mathbf{R}_k$ corresponding to the largest eigenvalues of the channel covariance matrix and reads

$$\mathbf{F}_k = [\mathbf{U}_k]_{:,1:T_{\mathrm{p}}} \mathbf{L}_k, \tag{6.5}$$

where $\mathbf{L}_k = \mathrm{diag}\big(\sqrt{\alpha_{k,1}}, \sqrt{\alpha_{k,2}}, \ldots, \sqrt{\alpha_{k,T_{\mathrm{p}}}}\big) \in \mathbb{R}^{T_{\mathrm{p}}\times T_{\mathrm{p}}}$ is a diagonal matrix that contains the weighted eigenvalues of $\mathbf{R}_k$. Specifically, a novel structure of the spatio-temporal common training matrix $\mathbf{S}_{\mathrm{p}} \in \mathbb{C}^{N\times T_{\mathrm{p}}}$ for the training-phase is designed based on the superposition of the weighted eigenvectors of $\mathbf{R}_k$, as expressed in (6.6),

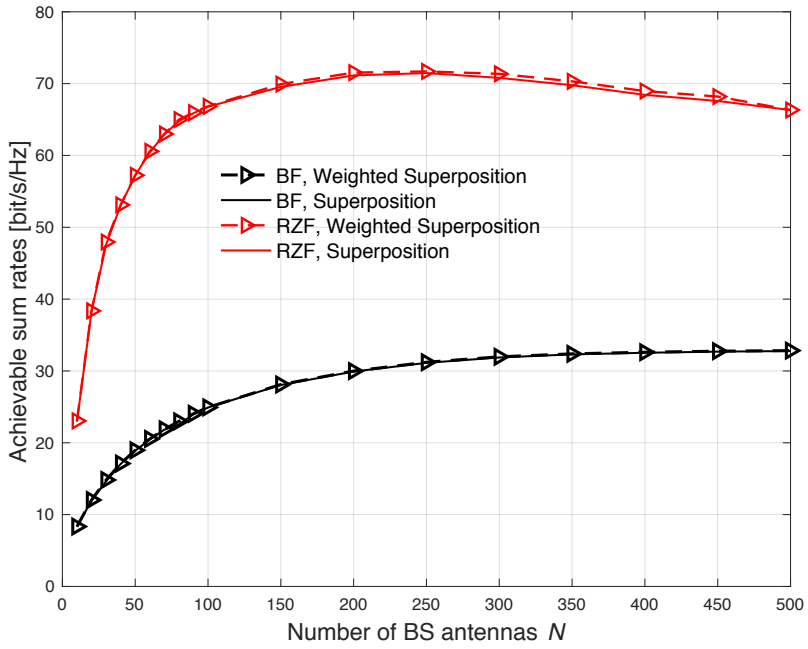


Figure 6.6 Achievable sum rate versus number of BS antennas N, comparing weighted and unweighted superposition training designs based on the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 10$ users.

$$\mathbf{S}_{\mathrm{p}} = \sum_{k=1}^{K} \frac{\mathbf{F}_k}{\left\| \sum_{k=1}^{K} \mathbf{F}_k \right\|_{\mathrm{F}}}, \tag{6.6}$$

where the pilot matrix in (6.6) is normalised based on the Frobenius norm to satisfy the power constraint $\mathrm{tr}\left(\mathbf{S}_{\mathrm{p}}^{\mathrm{H}}\mathbf{S}_{\mathrm{p}}\right) = T_{\mathrm{p}}$. The following section provides the numerical results, which characterise the performance of the BF and RZF precoders when the proposed weighted superposition training design is considered for DL channel estimation in the FDD massive MIMO systems.

6.3.1 Performance evaluation of the proposed weighted superposition training design for FDD massive MIMO systems

This subsection presents numerical results, which investigate the feasibility of the proposed weighted superposition training design in an FDD massive MIMO communication system when the users exhibit different spatial channel correlations. To this end, the achievable sum rate performance of the BF and RZF precoders based on the weighted superposition training design is compared with the conventional unweighted superposition design, which has been developed

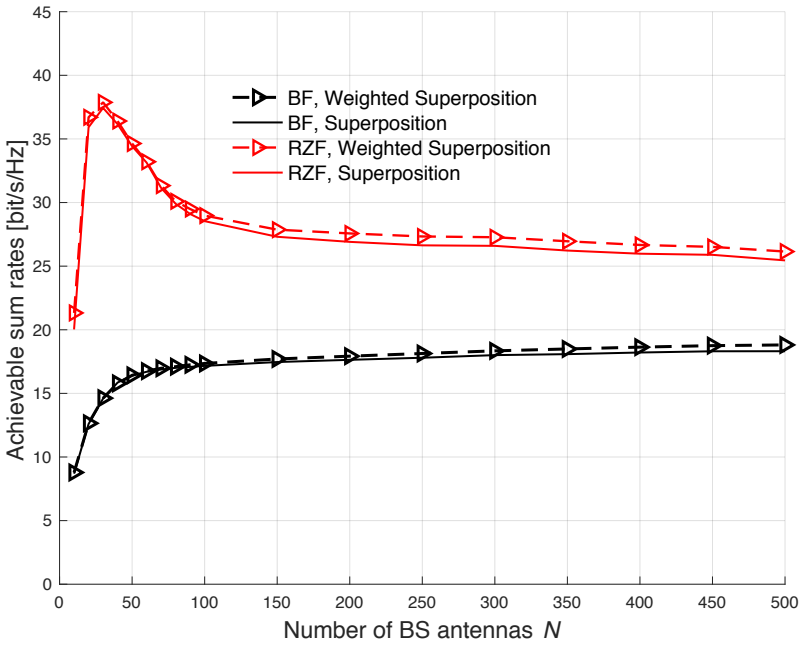


Figure 6.7 Achievable sum rate versus number of BS antennas N, comparing weighted and unweighted superposition training designs based on the OR model with $\omega_{\mathrm{H}} = 25°$, $D = 1$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 10$ users.

in Chapter 5. The numerical results are provided based on the scattering OR channel model when the salient system parameters K, N, T_{c}, SNR, AoAs, normalised antenna spacing D, and angular spread in the azimuth direction ω_{H} are used. In addition, a uniform linear array is considered in this subsection for the performance evaluation. Curves for the BF and RZF precoders are obtained numerically based on equations (5.27), (5.28), (5.29). These numerical analyses have been validated in Chapter 5 based on the Monte Carlo simulations, which provide an excellent agreement throughout. The solid lines denote the conventional (i.e. unweighted) superposition training design and the dash line with markers correspond to the newly proposed weighted superposition training design for both the BF and RZF precoders. The users' AoAs in the OR channel model are distributed in the range $[-60°, 60°]$ [53, 131], as provided in Chapter 3 and Chapter 5. For the OR model, parameters $\omega_{\mathrm{H}} = 5°$, $D = 1/2$ correspond to relatively strong correlation whereas $\omega_{\mathrm{H}} = 25°$, $D = 1$ imply relatively weak correlation.

Fig. 6.6 and Fig. 6.7 plot the achievable sum rate versus the number of BS antennas N, comparing the performance of the newly proposed weighted superposition training design in (6.6) with the conventional unweighted superposition training design in (5.25) for the DL channel estimation of an FDD massive MIMO system. The other salient system parameters in Fig. 6.6

and Fig. 6.7 were obtained with $T_c = 100$ symbols, SNR = 10 dB and $K = 10$ users, where the parameter values for the OR channel model are chosen based on different levels of correlations, i.e., $\omega_H = 5°$, $D = 1/2$ and $\omega_H = 25°$, $D = 1$, respectively. The results in Fig. 6.6 and Fig. 6.7 show that the proposed weighted superposition training sequence design for DL channel estimation achieves a marginal improvement in the sum rate performance compared with the rates obtained based on the unweighted superposition design.

The results in (Fig. 6.8 & Fig. 6.9) and (Fig. 6.10 & Fig. 6.11) investigate the achievable sum rate performance of the weighted superposition training sequence design when the number of users K is increased, i.e., $K = 12$ and $K = 14$, respectively, based on different levels of correlation, i.e., parameters $\omega_H = 5°$, $D = 1/2$ are considered in Fig. 6.8 & Fig. 6.10, whereas parameters $\omega_H = 25°$, $D = 1$ are considered in Fig. 6.9 & Fig. 6.11. The other salient system parameters remain unchanged at $T_c = 100$ symbols and SNR = 10 dB. The results in (Fig. 6.8 & Fig. 6.9) and (Fig. 6.10 & Fig. 6.11) demonstrate that the proposed weighted superposition training sequence design improves the sum rate performance marginally. Additionally, these results show that the peak sum rate performances obtained for both the BF and RZF precoders increase when the number of users K increases, as expected. The results in Fig. 6.7, Fig. 6.9, and Fig. 6.11 show that the achievable sum rate performance is degraded in the weak correlated channels, i.e., $\omega_H = 25°$, $D = 1$, as the number of transmit antennas N increases, as expected. Nonetheless, a feasible sum rate performances can be realised. Similar to Chapter 4, and Chapter 5, the achievable sum rate is enhanced by optimising the training sequence length through the maximisation of $(1 - T_p/T_c) \sum_{k=1}^{K} \log_2(1 + \text{SINR}_k)$, instead of only minimising the mean square error of the channel estimate, as typically considered in the conventional analyses of FDD and TDD systems. Overall, the results presented in this subsection demonstrate that proposed weighted superposition training design enhances the achievable sum rate performance of the FDD massive MIMO systems marginally compared with the conventional superposition design (i.e. unweighted), especially when the channels are relatively weak correlated.

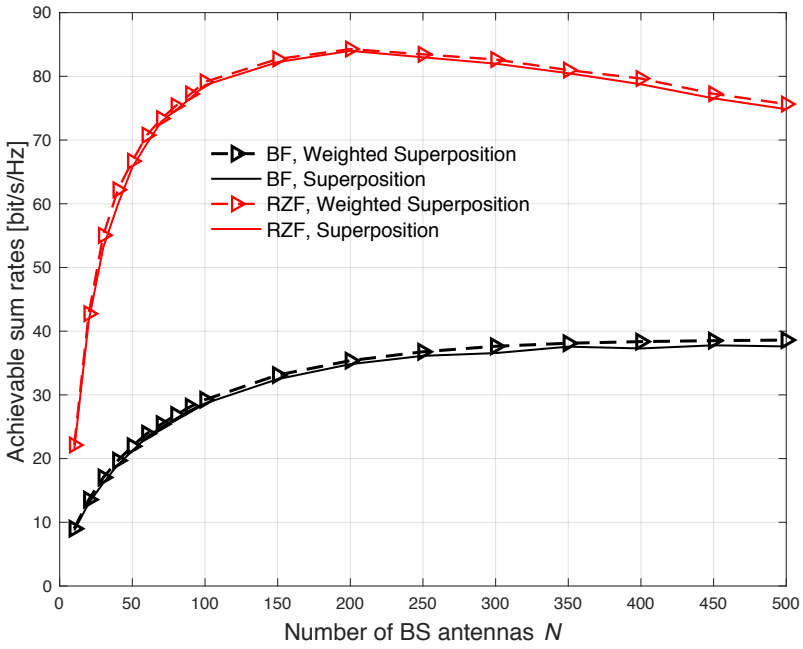


Figure 6.8 Achievable sum rate versus number of BS antennas N, comparing weighted and unweighted superposition training designs based on the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 12$ users.

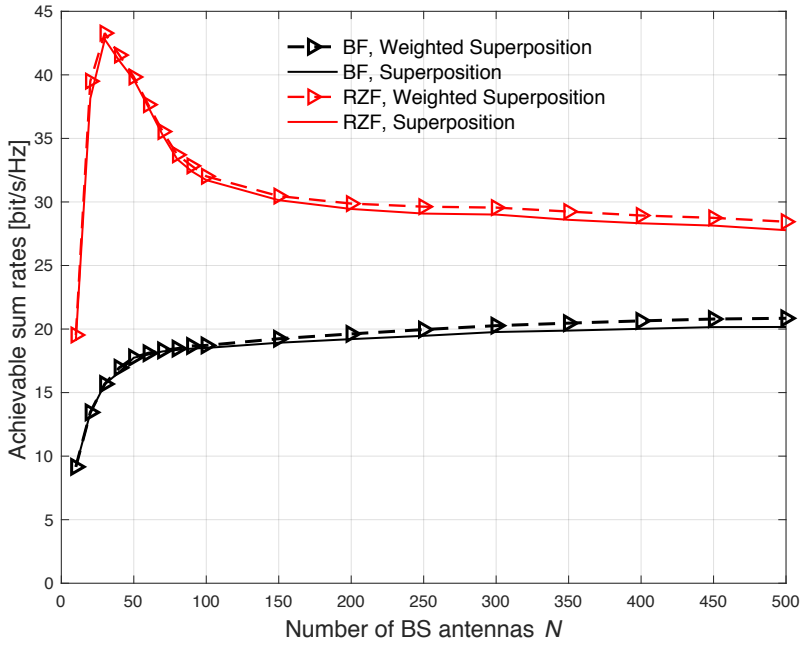


Figure 6.9 Achievable sum rate versus number of BS antennas N, comparing weighted and unweighted superposition training designs based on the OR model with $\omega_{\mathrm{H}} = 25°$, $D = 1$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 12$ users.

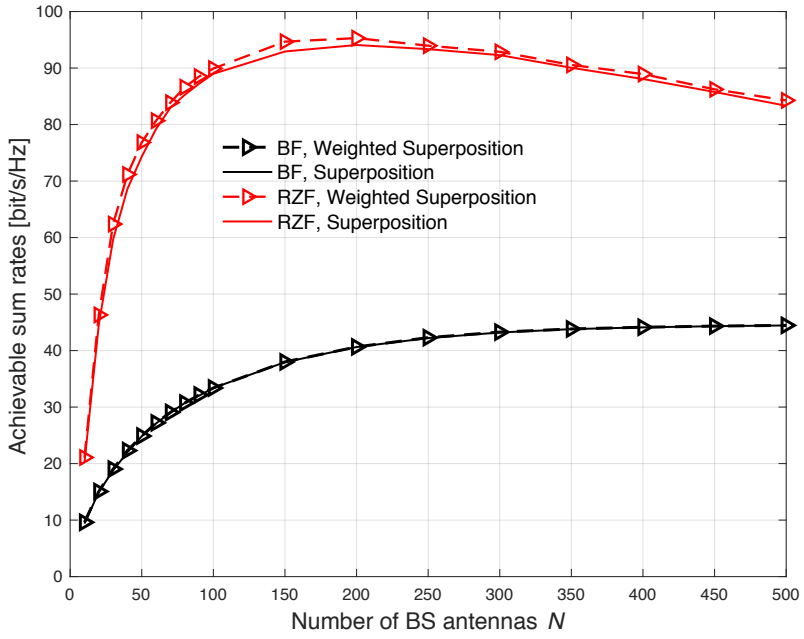


Figure 6.10 Achievable sum rate versus number of BS antennas N, comparing weighted and unweighted superposition training designs based on the OR model with $\omega_{\mathrm{H}} = 5°$, $D = 1/2$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 14$ users.

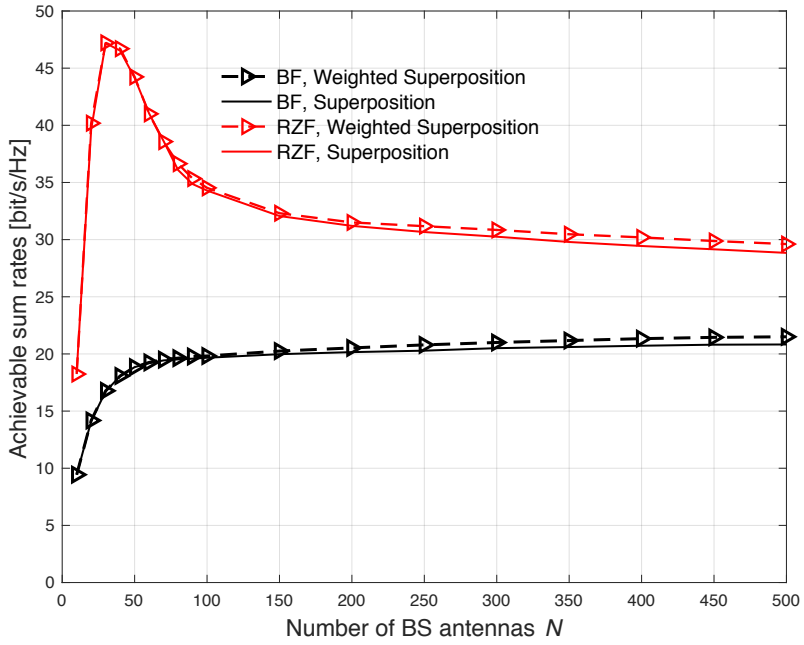


Figure 6.11 Achievable sum rate versus number of BS antennas N, comparing weighted and unweighted superposition training designs based on the OR model with $\omega_{\mathrm{H}} = 25°$, $D = 1$, $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB, and $K = 14$ users.

6.4 Training sequence design based on the linear superposition for uniform planar array in FDD massive MIMO systems

Although comprehensive investigation for the DL training sequence design and CSI estimation in the FDD massive MIMO systems has been carried out when a BS is equipped with a ULA, a DL pilot design for other practical antenna array configurations, such as the UPA, also needs to be investigated. To this end, a superposition training sequence design, which has been developed in Chapter 5, is adopted here for the DL CSI estimation of an FDD massive MIMO system when the UPA configuration is used at the BS. This also includes adopting the OR scattering channel model, described in Chapter 2 section 2.7.2, with the UPA configuration as described in [39, 42]. For a BS equipped with a ULA, the signals are propagated in the horizontal dimension (azimuth). This implies that the signals are radiated from directions within the same horizontal plane. For a BS equipped with a UPA, both the horizontal (azimuth) and vertical (elevation) dimensions are available, in which the signals can be radiated in both the horizontal and vertical directions.

In the following, the system geometric parameters of the channel covariance matrix $\mathbf{R}_k$ in the OR scattering channel model with UPA is presented. These parameters are determined based on the angular spread in the horizontal and vertical directions $\omega_{k,\mathrm{H}}$ and $\omega_{k,\mathrm{V}}$[1], angles of arrival in the horizontal and vertical directions $\theta_{k,\mathrm{H}}$ and $\theta_{k,\mathrm{V}}$, and normalised antenna spacing in the horizontal and vertical directions D_H and D_V in wavelengths, respectively. A uniform separation is assumed between the antenna elements of the BS array, i.e. a uniform space D is applied to both horizontal and vertical spaces; D_H, D_V, respectively.

Let $N = N_{\mathrm{H}}N_{\mathrm{V}}$ denote the total number of antenna elements at the BS with the UPA, where the horizontal and vertical channel covariance matrices are given as $\mathbf{R}_{k,\mathrm{H}} \in C^{N_{\mathrm{H}}\times N_{\mathrm{H}}}$ and $\mathbf{R}_{k,\mathrm{V}} \in C^{N_{\mathrm{V}}\times N_{\mathrm{V}}}$, respectively. Consequently, the (m,n)th element of $\mathbf{R}_{k,\mathrm{H}}$ and $\mathbf{R}_{k,\mathrm{V}}$ in the horizontal and vertical dimensions are given, respectively, in Toeplitz as [42]

[1]The subscripts H and V correspond to the horizontal and vertical directions, respectively.

$$[\mathbf{R}_{k,\mathrm{H}}]_{m,n} = \frac{1}{2\omega_{\mathrm{H}}} \int_{-\omega_{\mathrm{H}}+\theta_{k,\mathrm{H}}}^{\omega_{\mathrm{H}}+\theta_{k,\mathrm{H}}} e^{-j2\pi\, D\,(m-n)\sin(\Delta)} d\Delta, \tag{6.7}$$

$$[\mathbf{R}_{k,\mathrm{V}}]_{m,n} = \frac{1}{2\omega_{\mathrm{V}}} \int_{-\omega_{\mathrm{V}}+\theta_{k,\mathrm{V}}}^{\omega_{\mathrm{V}}+\theta_{k,\mathrm{V}}} e^{-j2\pi\, D\,(m-n)\sin(\Delta)} d\Delta, \tag{6.8}$$

The integration in (6.7) and (6.8) is computed numerically. The parameter ω_{V} represents the vertical angular spread and is given as [42]

$$\omega_{k,\mathrm{V}} = \frac{1}{2}\left(\arctan\left(\frac{d_k + s}{h}\right) - \arctan\left(\frac{d_k - s}{h}\right)\right). \tag{6.9}$$

The parameter $\theta_{k,\mathrm{V}}$ denotes the k-th UT AoA in the vertical direction and given as [42]

$$\theta_{k,\mathrm{V}} = \frac{1}{2}\left(\arctan\left(\frac{d_k + s}{h}\right) + \arctan\left(\frac{d_k - s}{h}\right)\right), \tag{6.10}$$

where d_k corresponds to the range where the UTs are located from the BS, h is the BS height in the elevation direction, and s is the radius of the scattering ring, all in metres. The narrow angular spread in the horizontal and vertical directions provides a low-rank structure for the channel covariance matrix, which implies, relatively, a strong spatial correlation. The channel covariance matrix $\mathbf{R}_k$ of the kth UT in the UPA configuration is obtained from the expression given in (6.11) [42],

$$\mathbf{R}_k = \mathbf{R}_{k,\mathrm{H}} \otimes \mathbf{R}_{k,\mathrm{V}}, \tag{6.11}$$

where $\otimes$ denotes the Kronecker operator. Based on the eigenvalue decomposition, described in Chapter 2, the k-th UT channel covariance matrix in the horizontal and vertical directions can be decomposed, respectively, as

$$\mathbf{R}_{k,\mathrm{H}} = \mathbf{U}_{k,\mathrm{H}} \mathbf{\Lambda}_{k,\mathrm{H}} \mathbf{U}_{k,\mathrm{H}}^{\mathrm{H}}, \tag{6.12}$$

$$\mathbf{R}_{k,\mathrm{V}} = \mathbf{U}_{k,\mathrm{V}} \mathbf{\Lambda}_{k,\mathrm{V}} \mathbf{U}_{k,\mathrm{V}}^{\mathrm{H}}, \tag{6.13}$$

Table 6.1 Uniform planar array configurations

N	N_{H}	N_{V}
16	4	4
32	8	4
64	8	8
128	16	8
256	16	16
512	32	16

where $\mathbf{U}_{k,\mathrm{H}} = [\mathbf{u}_{k,1}, \ldots, \mathbf{u}_{k,N_{\mathrm{H}}}] \in C^{N_{\mathrm{H}} \times N_{\mathrm{H}}}$ and $\mathbf{U}_{k,\mathrm{V}} = [\mathbf{u}_{k,1}, \ldots, \mathbf{u}_{k,N_{\mathrm{V}}}] \in C^{N_{\mathrm{V}} \times N_{\mathrm{V}}}$ are unitary matrices of the eigenvectors and $\mathbf{\Lambda}_{k,\mathrm{H}}$ and $\mathbf{\Lambda}_{k,\mathrm{V}}$ are a diagonal matrix of the eigenvalues of $\mathbf{R}_{k,\mathrm{H}}$ and $\mathbf{R}_{k,\mathrm{V}}$ arranged in descending order $\lambda_{k,1} \geq \lambda_{k,2} \geq \cdots \geq \lambda_{k,N_{\mathrm{H}}}$ and $\lambda_{k,1} \geq \lambda_{k,2} \geq \cdots \geq \lambda_{k,N_{\mathrm{V}}}$, respectively. A novel DL training sequence for the UPA is designed based on the linear superposition of sequences constructed from the channel covariance matrices in the horizontal (azimuth) and the vertical (elevation) directions by using the Kronecker operation. Specifically, the common spatio-temporal pilot matrix $\mathbf{S}_{\mathrm{p}} \in \mathbb{C}^{N \times T_{\mathrm{p}}}$ is constructed from the superposition of the first T_{p} eigenvectors of $\mathbf{R}_{k,\mathrm{H}}$ and $\mathbf{R}_{k,\mathrm{V}}$ that correspond to the largest eigenvalues of the channel covariance matrices in the horizontal and vertical directions, respectively, as expressed in (6.14).

$$\mathbf{S}_{\mathrm{p}} = \sum_{k=1}^{K} \mathbf{U}_{k,\mathrm{H}} \otimes \mathbf{U}_{k,\mathrm{V}} \tag{6.14}$$

The common pilot matrix in (6.14) is normalised by the Frobenius norm to satisfy the power constraint $\mathrm{tr}\left(\mathbf{S}_{\mathrm{p}}^{\mathrm{H}} \mathbf{S}_{\mathrm{p}}\right) = T_{\mathrm{p}}$. In the following subsection, the achievable sum rate performance of the proposed superposition training design for the UPA is investigated in the FDD multiuser massive MIMO communication systems with limited coherence time.

6.4.1 Performance evaluation of the proposed training sequence design for uniform planar array in FDD massive MIMO systems

In this subsection, the numerical results that characterise the achievable sum rate performance of the BF and RZF precoders based on the proposed superposition sequence design for

the UPA are presented. This includes providing a comparison with the state-of-the-art iterative algorithms for DL sequence designs in an FDD massive MIMO system, which have been already presented in Chapter 3. Here, the SCMI/SMSE iterative algorithms for the DL sequence optimisation in [53, 126] are implemented using the kth UT channel covariance matrix $\mathbf{R}_k$ with the UPA configuration, which is given by the expression in (6.11). Table 6.1 provides the number of antennas for the UPA configuration, which is used in the numerical evaluation.

Fig. 6.12 plots the achievable sum rate versus the number of BS antennas N, comparing the proposed superposition training design (6.14) with the SCMI/SMSE iterative algorithms for sequence designs based on the OR scattering channel model with the UPA for BF and RZF precoding. In particular, curves for the BF and RZF precoders are obtained numerically based on equations (5.27), (5.28), (5.29) where the users' channel covariance matrices in these expressions are replaced with the covariance matrix $\mathbf{R}_k$ of the kth UT obtained based on the UPA configuration in (6.11). The results in Fig. 6.12 are obtained with $T_{\mathrm{c}} = 100$ symbols, SNR = 10 dB and $K = 10$ users. The parameter values for the OR channel model with the UPA are chosen with an angular spread in the horizontal direction $\omega_{\mathrm{H}} = 5^{\circ}$, normalised antenna spacing $D = 1/2$, the mean AoAs in the horizontal direction are given in Chapter 3 and Chapter 5, the radius of the scattering ring $s = 30$ m, the BS height $h = 30$ m and the range of UTs distance from the BS is 200 m as given in [42]. As expected, results in Fig. 6.12 demonstrate that the RZF precoder achieves greater sum rate than the BF precoder for the massive MIMO system under consideration with the UPA. Furthermore, the results in Fig. 6.12 with the UPA show that the superposition training design achieves almost the same sum rate performance in comparison with the state-of-the-art SCMI/SMSE sequence designs. However, in the RZF precoder with the UPA configuration, the linear superposition design incurs a some loss in the rate performance in comparison with the rates obtained based on the state-of-the-art SCMI/SMSE iterative algorithms for sequence optimisation. Nonetheless, the results in Fig. 6.12 confirm that a feasible sum rate can be realised in an FDD massive MIMO system with UPA and limited coherence time using the superposition design, while reducing the computational complexity.

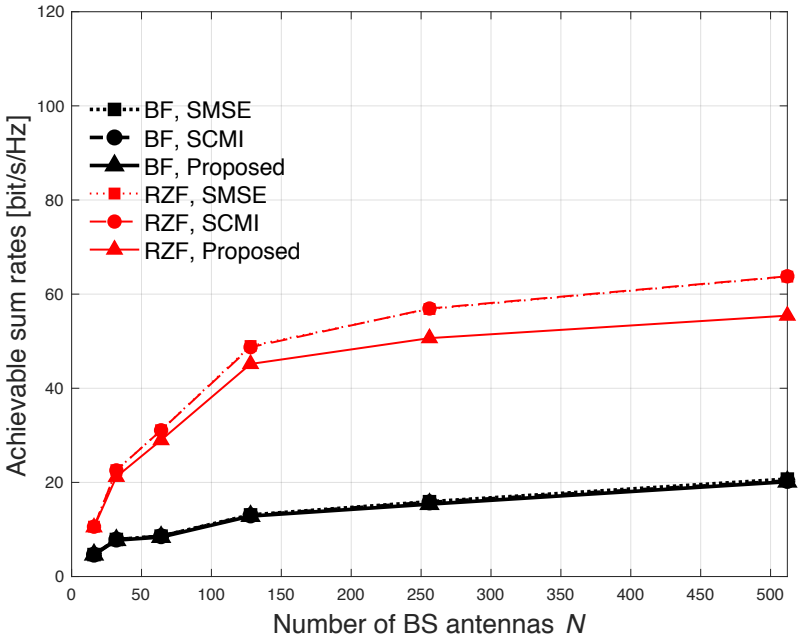


Figure 6.12 Achievable sum rate versus number of BS antennas N, $T_c = 100$ symbols, SNR = 10 dB, and $K = 10$ users, comparing different training sequence designs using the scattering OR channel model with UPA.

6.5 Summary

This chapter investigated further techniques to enhance the performance of DL training sequence design and improve the achievable sum rate of the FDD Massive MIMO systems with limited coherence time. Firstly, an investigation of the non-uniform power allocation was carried out between the training and data transmissions, which characterises the sum rate performance when the energy is optimised. Then, a new weighted superposition training design was developed to advance the conventional superposition training design in the FDD massive MIMO systems when users exhibit different spatial channel correlations. Additionally, the feasibility of the non-iterative superposition training design was investigated for the UPA.

The mathematical framework analyses of the SINR for the BF and RZF precoders, developed in Chapter 4 and Chapter 5 using the method of random matrix theory, are used in the performance evaluation of the proposed approaches in this chapter. This allows the numerical evaluation and the energy optimisation to be achieved with low-complexity. The analysis of the results in this chapter demonstrated that optimisation approaches with respect to choosing

an appropriate weighted energy in terms of the training and data transmissions, and consequently the weighted superposition design, achieved only a marginal enhancement in the sum rate performances of FDD massive MIMO systems. This suggests that such optimisation approaches are not necessary enough to warrant considering in the practical real-time systems due to the complexity that may add to the transceiver design. Nonetheless, the author of this thesis was motivated to characterise the DL achievable sum rate performance of an FDD massive MIMO system when such energy optimisation approaches are considered. Furthermore, the results showed that the superposition training design for the UPA provides comparable sum rate performance to the state-of-the-art iterative approaches when the BS employs BF precoding, while in the RZF precoder, the linear superposition design exhibits a some loss in the sum rate performance compared with the rates obtained by the iterative algorithms, especially for a large number of BS antennas (i.e. $N > 200$).

Chapter 7

Conclusions and Future Work

7.1 Conclusions

This thesis started with the aim of studying the transmission capabilities of an advanced physical layer technique known as a massive MIMO, which has been introduced as an essential part of future generation cellular networks in order to satisfy the rapidly increasing demand for the wireless data traffic. To this end, comprehensive research investigations of massive MIMO communications systems have been carried out throughout this thesis. The investigations included the presentation of the key potential gain of massive MIMO systems and the identification of technical challenges to their implementation.

This thesis demonstrated that having an accurate CSI estimate with single-stage precoding at the BS and limited coherence time is still a major challenge in massive MIMO communications systems. This thesis also explained that despite the attractiveness of the TDD mode of operation with massive MIMO when it comes to the CSI estimation, many currently deployed cellular networks are dominated by the FDD mode of operation. As such, this thesis has focused on investigating the performance of the FDD mode of operation in massive MIMO systems with single-stage precoding and limited coherence time. However, with the scaling up of the BS array in massive MIMO systems, the pilot overhead required for determining the CSI with single-stage precoding is overwhelming in FDD systems, thus obtaining a viable solution for the DL training sequence design becomes very challenging with limited coherence time. Therefore,

this thesis has particularly sought to explore feasible solutions with the lowest possible level of complexity and minimal overhead in order to overcome the technical challenges of designing a DL training sequence for the CSI estimation in the FDD massive MIMO communications systems.

Throughout this thesis, the focus has been on maximising the DL achievable sum rate performance. The sum rate is a composite variable, which depends on the opposing quantities $\log_2(1+\text{SINR}_k)$ and $(1-T_\text{p}/T_\text{c})$. Minimising the mean square error of the channel estimate maximises the $\log_2(1+\text{SINR}_k)$ term but it comes at the cost of the pre-log fraction term $(1-T_\text{p}/T_\text{c})$. Therefore, this thesis considered both terms to maximise the DL achievable sum rate for arbitrary values of BS antennas N and T_c. However, direct optimisation of the achievable sum rate is not a straightforward procedure, hence an asymptotic analytical method based on the random matrix theory was developed in this thesis. This allows a closed-form analysis to be achieved for advanced channel models, such as those considered in the majority of realistic communications systems, and thus, avoid executing computationally demanding Monte Carlo simulations. Two commonly prevailing types of linear precoders have been considered in this thesis, the BF precoder and the RZF precoder, which have been used in the majority of the open literature on MIMO system evaluation.

After providing an overview of massive MIMO communications systems, the investigations began by reviewing the state-of-art research undertaken to tackle the challenges faced by the FDD mode of operation in massive MIMO systems, which motivated the requirement for the new contributions of this thesis. Chapter 3 looked at gradient-based iterative algorithms in order to optimise the DL training sequences and identify the best sum rate performances that can be achieved in FDD massive MIMO systems with single-stage precoding when the users exhibit distinct spatial correlations. A conjugate gradient descent iterative algorithm was proposed, which has the ability to provide a robust sum rate performance and enhance the convergence rate compared with the state-of-the-art iterative algorithms. However, the analyses of the computational complexity of different gradient-based optimisation approaches in Chapter 3 showed that iterative solutions for designing the DL training sequences in the FDD massive MIMO systems require hundreds or thousands of iterations to reach convergence, which is still overly complex, thus constraining the approach. In particular, utilising such iterative approaches for

the DL training sequence design could be a bottleneck in real-time systems where the CSI must be estimated more frequently to track the user channels variation, especially when a scenario with limited coherence time interval is considered. Therefore, there was a crucial requirement to develop a non-iterative solution for designing a viable DL training sequence, which can be more feasible, and thus, satisfies the practical implementation concerns. Furthermore, such iterative optimisation approaches provide no closed-form solutions on either the optimum structure of a DL training sequence, nor on the minimum pilot length that maximises the achievable sum rate under a limited coherence time.

As such, Chapter 4 investigated a non-iterative approach for the DL training sequence design in an FDD massive MIMO system based on a scenario where users exhibit common spatial correlations. The focus was on providing a tractable framework analysis for the optimum training structure and deriving the optimum pilot sequence length that maximises the achievable sum rate with limited channel coherence time. To this end, a computationally feasible solution for the DL training sequence was developed and a tractable framework analysis for the optimum DL training structure was produced. Specifically, Chapter 4 showed explicitly how to assign an optimum unique training sequence to each of the users individually, whereby the eigenvectors based on the transmit covariance matrix that correspond to the largest eigenvalues of any user was the same for all users. Additionally, a new analytically closed-form solution for the optimum DL training sequence length of the BF precoder was provided. The computational complexity of the non-iterative training sequence design was analysed and compared with the state-of-the-art iterative algorithm for sequence designs. The results showed that the non-iterative approach for DL training design achieved a significant reduction in the computational complexity compared with the sequence designs based on iterative algorithms.

However, designing a viable DL training scheme with reduced complexity in a more realistic propagation conditions where users exhibit distinct spatial correlations, and which is feasible with single-stage precoding and limited coherence time, remains a challenge in FDD massive MIMO systems. Further, finding the optimum training sequence length that maximises the achievable sum rate in FDD massive MIMO systems with different correlation patterns is very challenging due to the technical challenge of deriving closed-form solutions based on such channel conditions. To address these challenges, Chapter 5 built on the tractable frame-

work analysis developed in Chapter 4 . To this end, the principle of linear superposition of sequences constructed from the eigenvectors of distinct users' channel correlation matrices was developed with single-stage precoding. Based on the superimposed training design, a new analytical closed-form solution for the optimum training sequence length, which maximises the DL achievable sum rate in the BF precoder with distinct spatial correlations between users, was developed. The analysis of the results showed that the optimum training sequence length that was analytically optimised for the BF precoder with the P-DoF channel model was sufficient to predict the achievable sum rate performance of the RZF precoder, which remains near optimal. This observation also remained valid with the more practical OR scattering channel model with only a small loss in performance. In addition, the results showed that the superposition training design provided comparable sum rate performance to the state-of-the-art sequence designs based on iterative algorithms. The computational complexity of the superposition training sequence design was quantified and compared with the state-of-the-art iterative approaches. The analyses of the complexity results in Chapter 5 showed that more than four orders-of-magnitude reduction in the computational complexity was achieved using the superposition training sequence design, which signifies the feasibility of this approach for practical implementations compared with state-of-the-art iterative algorithms. Furthermore, an analytical asymptotic approximation that characterises the achievable sum rate performance of the BF and RZF precoders has been developed using large dimension random matrix theory and the P-DoF channel model. The numerical results in this thesis showed that these asymptotic analyses were accurate for practical, finite systems parameters in terms of N and K.

The analyses of the results have shown that the diversity of spatial channel correlations between multiple users significantly enhances the achievable sum rate of an FDD massive MIMO system using DL channel estimation. Furthermore, comparison between the correlated channels with K independent channel covariance matrices and uncorrelated channels with identical channel covariance matrices was also provided. The results showed that even when the channels are uncorrelated, a feasible sum rate performance with a DL CSI estimation was realised in an FDD massive MIMO system. This can be justified by the fact that the optimum training sequence length was obtained by maximising $(1 - T_{\mathrm{p}}/T_{\mathrm{c}}) \sum_{k=1}^{K} \log_2(1 + \mathrm{SINR}_k)$ instead of minimising the mean square error of the channel estimate only, as typically considered in the conventional

FDD and TDD systems.

Comparison between the sum rate of perfect and imperfect knowledge of the CSI estimation was provided, which allows the effect of DL training sequences on the achievable sum rate of BF and RZF precoding in an FDD massive MIMO system to be explored. The results indicated that in FDD systems with limited coherence time, the DL training sequence length saturates and the achievable sum rate decreased as N asymptotically increased due to imperfect DL CSI estimation. This thesis also provided comparisons between the achievable sum rates of the UL CSI estimation as conventionally used in a TDD system and the DL CSI estimation in an FDD system. The results showed that for practical BS array sizes of $N < 250$ antennas, the sum rate of a massive MIMO system using DL channel estimation is comparable to the performance of a TDD system in moderately to strongly correlated channels.

To improve the performance of the non-iterative DL training approaches, Chapter 6 investigated a non-uniform power allocation between the training and data transmissions. Chapter 6 also looked at enhancing the performance of the linear superposition training design developed in Chapter 5 by proposing a new weighted superposition training design in the FDD massive MIMO systems when users exhibit different spatial correlations. In addition, Chapter 6 studied the feasibility of the superposition training design for other practical antenna array configurations, such as the UPA. Results characterising the achievable sum rate performance for the BF and RZF precoders in the P-DoF and the OR channel models were provided. The results showed that optimisation approaches with respect to choosing an appropriate weighted energy in terms of the training and data transmissions and the weighted superposition design achieved a marginal improvement in the sum rate performances of FDD massive MIMO systems, which could not always justify the additional complexity to implement the schemes. Finally, the analysis of the results indicated that when the BS employs the UPA, the superposition training design offers comparable sum rate performance to the state-of-the-art iterative approaches for the DL sequence designs, while again considerably reducing the computational complexity.

Overall, the findings in this thesis were supported by rigorous mathematical analyses, which tightly agree with the simulated results, which underpin the contributions of this research. This thesis has shown that a feasible sum rate performance in a massive MIMO system can be achieved with a significant reduction in the computational complexity when finite length

training sequences are used in the DL CSI estimation of a non-reciprocal wireless channel as in FDD based systems.

7.2 Future work

This subsection suggests some possible potential scope for future lines of research.

- **The consideration of Rician fading channels**

  The vast majority of massive MIMO research has assumed that the channels are either modelled with uncorrelated Rayleigh fading channels based on a rich scattering environment or with spatially correlated Rayleigh fading channels based on a limited scattering environment. However, since one of the applications of massive MIMO system is in a mm-wave band, where the line-of-sight (LoS) effect may dominate [134], investigating the FDD massive MIMO systems in more general fading scenarios, which takes into account the LoS conditions in addition to the NLoS components, is an interesting research topic that should be investigated in the future.

- **The extension to multi-cell networks**

  Although a comprehensive investigation of DL pilot design in an FDD massive MIMO system based on a single-cell scenario has been carried out in this thesis, the training sequence assignment in multi-cell systems require a thorough investigation and evaluation. This includes providing an analytical closed-form solution that characterises the achievable sum rate for the BF and RZF precoders. Furthermore, the future generation of wireless networks are also expected to remove the cell boundaries, so that applying cell-free networks [135–138]. Therefore, the DL training sequence design and CSI estimation in an FDD system with cell-free networks is also worth investigating in the future.

- **Other types of antenna array configurations**

  In this thesis, the DL training sequence has been evaluated for the massive MIMO system when the BS is equipped with a ULA or UPA. The achievable sum rate performance evaluation of other types of antenna array configurations, such as non-uniform, cylindrical,

and circular array configurations should also be investigated in the future [139, 140]. One would expect that with other antennas array configurations, the degrees of correlations between the antenna elements may be varied, and thus, the achievable sum rate performance may be affected accordingly.

- **Uplink Feedback design with FDD massive MIMO systems**

  Although this thesis concentrated entirely on designing feasible solutions for the DL training sequence design and minimising the training duration over a limited coherence time, optimising the achievable sum rate with respect to uplink feedback is another potential research direction that could be addressed in future work. The presence of spatial correlations could be beneficial in the uplink feedback design in the FDD massive MIMO systems. The author is particularly interested in studying the uplink feedback and investigating how it would affect the DL achievable sum rate performance when the channel coherence time remains finite.

- **Investigate the massive MIMO system performance with very different second order channel statistics**

  Throughout this thesis, the achievable sum rate performance of a massive MIMO system has been characterised in the realistic channel model, i.e., OR channel model, when all users exhibit either full-rank (uncorrelated) or rank-deficient covariance matrices. The author is also interested in investigating the sum rate performance in the more general scenario where some of the users exhibit full-rank (uncorrelated) covariance matrices while other users exhibit rank-deficient covariance matrices.

- **Investigate alternative analytical channel models**

  Investigate alternative analytical channel models to the P-DoF model, which, while effective, is overly simplistic and most necessarily accurate.

Appendix A

This Appendix briefly explains the methodology for calculating the flops per iteration for the different iterative algorithms, presented in Chapter 3, Section 3.5.

- The following explains how to calculate the total number of multiplications and additions in flops of a matrix-matrix product of an $N \times N$ matrix with an $N \times N$ matrix. Let $\mathbf{A} = \mathbf{BC}$, where $\mathbf{B}$ and $\mathbf{C}$ are two matrices each with a size of $N \times N$. Each element in the obtained matrix $\mathbf{A}$ involves an N multiplications and $N - 1$ additions. This is valid for any matrix, which depends on the number of rows and columns of the matrix. Hence, the first element of the obtained matrix $\mathbf{A}$ involves an (N) multiplication + $(N-1)$ addition = $2N - 1$ flops. The whole matrix-matrix multiplication operation then requires $(2N - 1)(N)(N) = 2N^3 - N^2$ flops. The same procedure holds for all the matrix-matrix products, which depends on the size of each matrix considered.

- In the following, the methodology for obtaining the total number of flops of the gradient term that is involved in both the proposed CGD and the SMSE algorithms, and which is given by $\mathbf{R}_k^2\mathbf{S}_\mathrm{p}\left(\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{R}_k\mathbf{S}_\mathrm{p} + \mathbf{I}_{T_\mathrm{p}}\right)^{-1} + \mathbf{R}_k\mathbf{S}_\mathrm{p} \times \left(\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{R}_k\mathbf{S}_\mathrm{p} + \mathbf{I}_{T_\mathrm{p}}\right)^{-1}\mathbf{S}_\mathrm{p}^\mathrm{H}\,\mathbf{R}_k^2\mathbf{S}_\mathrm{p}\left(\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{R}_k\mathbf{S}_\mathrm{p} + \mathbf{I}_{T_\mathrm{p}}\right)^{-1}$, is provided. The calculation involves counting different sizes of a matrix-matrix product, matrix-matrix addition, and a squared matrix inversion. The number of flops required in these different operations have already been given in Chapter 3 section 3.5. Starting the count of the flops operations from the right hand side of the term given above, the following flops operations are required $\Big(T_\mathrm{p}^2 + 2N^2T_\mathrm{p} - NT_\mathrm{p} + 2NT_\mathrm{p}^2 - T_\mathrm{p}^2 + T_\mathrm{p}^3 + 2NT_\mathrm{p}^2 - NT_\mathrm{p} + 2N^2T_\mathrm{p} - NT_\mathrm{p} + 2N^2T_\mathrm{p} - NT_\mathrm{p} + 2NT_\mathrm{p}^2 - T_\mathrm{p}^2 + T_\mathrm{p}^3 + 2T_\mathrm{p}^3 - T_\mathrm{p}^2 + T_\mathrm{p}^2 + 2N^2T_\mathrm{p} - NT_\mathrm{p} + 2NT_\mathrm{p}^2 - NT_\mathrm{p} + 2N^2T_\mathrm{p} - NT_\mathrm{p} + NT_\mathrm{p} + T_\mathrm{p}^3 + T_\mathrm{p}^2 + 2N^2T_\mathrm{p} - NT_\mathrm{p} + 2NT_\mathrm{p}^2 - T_\mathrm{p}^2 + 2NT_\mathrm{p}^2 - NT_\mathrm{p} +$

$2N^2T_\mathrm{p} - NT_\mathrm{p} + 2N^2T_\mathrm{p} - NT_\mathrm{p}\Big)$. Combining all the terms above leads directly to the total number of flops required, which is given as $\big(5T_\mathrm{p}^3 + 16N^2T_\mathrm{p} + 14NT_\mathrm{p}^2 - 2T_\mathrm{p}^2 - 10NT_\mathrm{p}\big)$.

- In order to calculate the squared Frobenius norm, also known as the squared Hilbert Schmidt norm [123], which is denoted by $||\cdot||_\mathrm{F}$, of an $N\times T_\mathrm{p}$ matrix requires the following summations and multiplications. It follows from summing up an $N\times T_\mathrm{p}$ entries and $N\times T_\mathrm{p}$ multiplications. Combining these two operations together, yields the total number of flops of $2NT_\mathrm{p} - 1$.

- In the following, the methodology for calculating the computational complex of the SCMI iterative algorithm is described. This complex calculation is involved counting different sizes of a matrix-matrix product, matrix-matrix addition, and a squared matrix inversion. Chapter 3 section 3.5 provides the number of flops required in these matrices operations. Following the same procedure described above, calculating the term $\sum_{k=1}^{K}\mathbf{R}_k\mathbf{S}_\mathrm{p}\big(\mathbf{S}_\mathrm{p}^\mathrm{H}\mathbf{R}_k\mathbf{S}_\mathrm{p} + \mathbf{I}_{T_\mathrm{p}}\big)^{-1}$. This term is a first step in the SCMI algorithm which requires $(K-1)\big(T_\mathrm{p}^3 + 4NT_\mathrm{p}^2 + 4N^2T_\mathrm{p} - 3NT_\mathrm{p}\big)$ flops.

- Steps 2 and 3 in the SCMI algorithm requires $T_\mathrm{p}^3 - T_\mathrm{p}^2 + 2NT_\mathrm{p}^2 + 4NT_\mathrm{p} - 1$ flops.

- Combining all the described above flops operations yields, $\big(\, KT_\mathrm{p}^3 + 4NT_\mathrm{p}^2K + 4N^2T_\mathrm{p}K - 3NT_\mathrm{p}K - T_\mathrm{p}^3 - 4NT_\mathrm{p}^2 - 4N^2T_\mathrm{p} + 3NT_\mathrm{p} + T_\mathrm{p}^3 + 2T_\mathrm{p}^2N - T_\mathrm{p}^2 + 4NT_\mathrm{p} - 1\,\big)$. This can be simplifies to $\big(\, KT_\mathrm{p}^4 + 4NT_\mathrm{p}^2K + 4N^2T_\mathrm{p}K - 3NT_\mathrm{p}K - 2T_\mathrm{p}^2N - T_\mathrm{p}^2 - 4N^2T_\mathrm{p} + 7NT_\mathrm{p} - 1\big)$ flops. This also can further be simplified to $\big(\, KT_\mathrm{p}^3 + T_\mathrm{p}^2(N(4K-2)-1) + NT_\mathrm{p}(4N(K-1)+7-3K) - 1\,\big)$ flops. Finally, the total complexity of the SCMI algorithm in flops is given as $\big(\, T_\mathrm{p}^3K + T_\mathrm{p}^2\big(N\,(4K-2)-1)\big) + 4N^2T_\mathrm{p}(K-1) + NT_\mathrm{p}\,(7-3K) - 1\,\big)$.

- In the following, the methodology for calculating the computational complexity of the SMSE iterative algorithm is briefly explained. Step 1 in the SMSE iterative algorithm is to calculate the gradient, which similar to the procedure described above. Step 2 in the SMSE iterative algorithm requires $4N^2T_\mathrm{p}+NT_\mathrm{p}$ flops. Step 3 in the SMSE requires $2NT_\mathrm{p}-1$ flops. Steps 4 and 5 require $(t_d + 1)\,23N^2T_\mathrm{p}$, and $(t_h + 2)\,23N^2T_\mathrm{p}$, respectively. Step 6 requires $23N^2T_\mathrm{p}$. Combining all the flops leads to the total number of computational complexity of the SMSE iterative algorithm.

Appendix B

In the following, a proof of the asymptotically tight approximation of the SINR for BF in Theorem 4.1 is provided. Recalling the analytical P-DoF channel model given in (2.6), and due to the condition of common spatial channel correlation considered in Chapter 4, the covariance matrices of the users are all the same and of the form $\mathbf{R} = \frac{N}{P}\mathbf{A}\mathbf{A}^{\mathrm{H}}$, where $\mathbf{A} \in \mathbb{C}^{N\times P}$ is constructed from $P \leq N$ columns of an arbitrary $N \times N$ unitary matrix so that $\mathbf{A}^{\mathrm{H}}\mathbf{A} = \mathbf{I}_P$. As explained in Chapter 2, the trace operator of a square matrix $\mathbf{R}$, denoted by $\mathrm{tr}(\mathbf{R})$, equals to the sum of diagonal entries (eigenvalues) of $\mathbf{R}$. Also, the $\mathrm{tr}(\mathbf{A}^{\mathrm{H}}\mathbf{A}) = \mathrm{tr}(\mathbf{A}\mathbf{A}^{\mathrm{H}})$. Additionally, $\mathrm{tr}(\mathbf{R}^2) = \frac{N^2}{P^2}P$. Substituting these expressions with the P-DoF into (4.3), provides the expression in (B.1) as

$$\overline{\mathrm{SINR}}^{\mathrm{BF,PF}} = \frac{N}{\frac{1}{\rho_{\mathrm{d}}} + \frac{K}{N}\left(\frac{N^2}{P^2}\right)P}, \tag{B.1}$$

which leads directly to (4.3). At high SNR case, so that when $\rho_{\mathrm{d}} \to \infty$, the left hand side in the denominator of (4.3) approaches zero. Thus, expression in (4.3) is simplified to

$$\overline{\mathrm{SINR}}^{\mathrm{BF,PF}} = \frac{N}{\frac{KN}{P}}. \tag{B.2}$$

The expression in (B.2) can be further simplified to

$$\overline{\mathrm{SINR}}^{\mathrm{BF,PF}} = \frac{P}{K}. \tag{B.3}$$

Finally, the expression (B.3) is substituted into the expression (4.2), which yields (4.15), completing the proof.

Appendix C

This Appendix explains how to find a closed-form solution to the fixed point algorithm. Let the following simplified expression,

$$X^t = \frac{A}{\frac{B}{1+X^{(t-1)}} + C}, \tag{C.1}$$

where A, B, C could be any auxiliary variable. Considering the limit $t \to \infty$, expression (C.1) can be written as

$$X = \frac{A}{\frac{B}{1+X} + C}. \tag{C.2}$$

Then, to find a unique solution to the variable X in the expression (C.2), let the equality in following expression

$$CX^2 + (B + C - A)X - A = 0. \tag{C.3}$$

Hence, a unique positive solution to the variable X can be obtained as in (C.4).

$$X = \frac{A - B - C + \sqrt{C^2 + (2\,B + 2\,A)\,C + B^2 - 2AB + A^2}}{2C} \tag{C.4}$$

The closed-form expression in (C.4) has been used in the RZF precoder simplification.

Appendix D

This Appendix explains briefly how the expression for the SINR of RZF precoding (4.16) is obtained. Unlike the expression of the BF precoder, the SINR approximation for the RZF precoder is provided in terms of several auxiliary variables, which arise from the asymptotic RMT analysis. Nonetheless, the SINR expression of the RZF precoder under a general correlation model with perfect CSI estimation in (4.5) depends on $\mathbf{R}$ and not the covariance matrix of the MMSE channel estimation $\mathbf{\Phi}$.

Using the analytical P-DoF channel model given in (2.6) with $\mathbf{R} = \frac{N}{P}\mathbf{A}\mathbf{A}^{\mathrm{H}}$, where $\mathbf{A} \in \mathbb{C}^{N\times P}$ is constructed from $P \leq N$ columns of an arbitrary $N \times N$ unitary matrix so that $\mathbf{A}^{\mathrm{H}}\mathbf{A} = \mathbf{I}_P$. As explained in Appendix B, the trace operator of a square matrix $\mathbf{R}$, denoted by $\mathrm{tr}(\mathbf{R})$, equals to the sum of diagonal entries (eigenvalues) of $\mathbf{R}$. Further, the $\mathrm{tr}(\mathbf{A}^{\mathrm{H}}\mathbf{A}) = \mathrm{tr}(\mathbf{A}\mathbf{A}^{\mathrm{H}})$. Consequently, $\mathrm{tr}(\mathbf{R}) = \frac{N}{P}P$. Substituting aforementioned expressions with the P-DoF into (4.5), and using the closed-form solution of the fixed point algorithm explained in Appendix C, a unique solution in expression (4.17) is obtained. Simpler procedure is used to obtain the expressions in (4.18) and (4.19). A computer program, Wolfram Mathematica, was used to aid with the analysis simplification of the RZF precoder.

Appendix E

This Appendix provides a proof of the expressions in (4.28) and (4.29). As explained in Chapter 2, the unitary of an $N \times N$ matrix i.e. $\mathbf{U} \in \mathbb{C}^{N \times N}$ of the eigenvectors is orthogonal and satisfied $\mathbf{U}^{\mathrm{H}}\mathbf{U} = \mathbf{I}_N$. It also showed that if $\mathbf{U} \in \mathbb{C}^{N \times T_{\mathrm{p}}}$, then $\mathbf{U}^{\mathrm{H}}\mathbf{U} = \mathbf{I}_{T_{\mathrm{p}}}$. Consider the pilot matrix in (E.1)

$$\mathbf{S}_{\mathrm{p}} = \mathbf{U}_{T_{\mathrm{p}}} = \left[\mathbf{u}_1, \ldots, \mathbf{u}_{T_{\mathrm{p}}}\right]. \tag{E.1}$$

Using the unitary matrix properties explained above, the EVD in (4.22), and the pilot matrix in (E.1), the covariance matrix of the MMSE channel estimation $\mathbf{\Phi} \in \mathbb{C}^{N \times N}$ in (4.21) can be simplified as

$$\mathbf{\Phi} = \rho_{\mathrm{p}}\mathbf{U}_{T_{\mathrm{p}}}\mathbf{\Lambda}_{T_{\mathrm{p}}}^2\left(\rho_{\mathrm{p}}\mathbf{\Lambda}_{T_{\mathrm{p}}} + \mathbf{I}_{T_{\mathrm{p}}}\right)^{-1}\mathbf{U}_{T_{\mathrm{p}}}^{\mathrm{H}}, \tag{E.2}$$

where $\mathbf{\Lambda}_{T_{\mathrm{p}}} \in \mathbb{R}^{T_{\mathrm{p}} \times T_{\mathrm{p}}}$ is a diagonal matrix with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{T_{\mathrm{p}}}$. As explained in Chapter 2, the trace operator of a square matrix equals to the sum of diagonal entries (eigenvalues) of that matrix. In addition, the trace is a linear mathematical operation, which is invariant to unitary rotation and a change of basis, so that $\mathrm{tr}(\mathbf{U}^{\mathrm{H}}\mathbf{U}) = \mathrm{tr}(\mathbf{U}\mathbf{U}^{\mathrm{H}})$. Thus, taking the trace of the covariance matrix of the MMSE channel estimation in (E.2), which is a square matrix, leads to

$$\mathrm{tr}(\mathbf{\Phi}) = \sum_{n=1}^{T_{\mathrm{p}}} \frac{\rho_{\mathrm{p}}\lambda_n^2}{\rho_{\mathrm{p}}\lambda_n + 1}, \tag{E.3}$$

where $\mathbf{U}^{\mathrm{H}}\mathbf{U} = \mathbf{I}_{T_{\mathrm{p}}}$. Finally, dividing both the nominator and denominator by ρ_{p}, directly leads to the expression in (4.28), completing the proof. Using the EVD in (4.22) and the trace mathematical operation, the expression in (4.29) can be directly obtained, following the same procedure

of the proof described above, which can be also written as in (E.4).

$$\mathrm{tr}(\mathbf{R}\boldsymbol{\Phi}) = \sum_{n=1}^{T_\mathrm{p}} \frac{\rho_\mathrm{p}\lambda_n^3}{\rho_\mathrm{p}\lambda_n + 1}. \tag{E.4}$$

This completing the proof.

Appendix F

This Appendix explains how to obtain a closed-form analysis for the SINR of the BF precoder based on the P DoF channel model. Recalling the P-DoF channel model defined in (2.6), and due to the condition of common correlation for all the users, the covariance matrices of all the users are the same and of the form $\mathbf{R} = \frac{N}{P}\mathbf{A}\mathbf{A}^{\mathrm{H}}$, where $\mathbf{A} \in \mathbb{C}^{N\times P}$ is constructed from $P \leq N$ columns of an arbitrary $N \times N$ unitary matrix so that $\mathbf{A}^{\mathrm{H}}\mathbf{A} = \mathbf{I}_P$.

Recalling the EVD of the transmit channel covariance matrix $\mathbf{R}$, defined in (2.4), where the eigenvalues $\mathbf{\Lambda}$ of $\mathbf{R}$ with P-DoF channel model is now a diagonal matrix containing the ordered eigenvalues of $\mathbf{R}$, namely, $\lambda_1 = \lambda_2 = \cdots = \lambda_P = N/P$ and $\lambda_{P+1} = \cdots = \lambda_N = 0$. As described in Appendix E, the covariance matrix of the MMSE channel estimation in (E.2), can be written as

$$\mathrm{tr}(\mathbf{\Phi}) = \sum_{n=1}^{T_{\mathrm{p}}} \frac{\rho_{\mathrm{p}}\lambda_n^2}{\rho_{\mathrm{p}}\lambda_n + 1}, \tag{F.1}$$

Since the eigenvalues are known in the P-DoF channel model, this simplifies the channel estimate covariance matrix in (F.1) to

$$\mathrm{tr}(\mathbf{\Phi}) = \frac{\rho_{\mathrm{p}}(N^2/P^2)\,T_{\mathrm{p}}}{\rho_{\mathrm{p}}\,(N/P) + 1}. \tag{F.2}$$

Multiplying both the nominator and denominator by (P^2), yields,

$$\mathrm{tr}(\mathbf{\Phi}) = \frac{\rho_{\mathrm{p}}(N^2)\,T_{\mathrm{p}}}{\rho_{\mathrm{p}}\,(NP) + P^2}. \tag{F.3}$$

By dividing both the nominator and denominator with N^2, leads to

$$\mathrm{tr}(\mathbf{\Phi}) = \frac{\rho_{\mathrm{p}}\, T_{\mathrm{p}}}{\rho_{\mathrm{p}}\,(P/N) + P^2/N^2}. \tag{F.4}$$

The expression in (F.4), can also be rewritten as

$$\mathrm{tr}(\mathbf{\Phi}) = \frac{\rho_{\mathrm{p}}\, T_{\mathrm{p}}}{(\rho_{\mathrm{p}} + P/N)\, P/N}. \tag{F.5}$$

Recalling the SINR expression for the BF precoder under a general channel correlation model, which is given in (F.6)

$$\overline{\mathrm{SINR}}^{\mathrm{BF}} = \frac{\big(\mathrm{tr}(\mathbf{\Phi})\big)^2}{\frac{1}{\rho_{\mathrm{d}}}\mathrm{tr}(\mathbf{\Phi}) + K\,\mathrm{tr}(\mathbf{R}\mathbf{\Phi})}, \tag{F.6}$$

where $\mathbf{\Phi}$ is the covariance matrix of the MMSE channel estimate. The following finalises the proof of the closed-form analysis of the SINR for the BF precoder with the P-DoF channel model. Since the eigenvalues of the P-DoF channel model are known in, a simplified expression to the channel estimate covariance matrix in (F.1) can be provided as in (F.2). Consequently, expression in (E.4) can also be written as

$$\mathrm{tr}(\mathbf{R}\mathbf{\Phi}) = \frac{\rho_{\mathrm{p}}(N^3/P^3)\, T_{\mathrm{p}}}{\rho_{\mathrm{p}}(N/P) + 1}, \tag{F.7}$$

Let assume that $N/P = C$ for notational convenience, and hence, the expressions in (F.2) and (F.7) can be written, consequently, as

$$\mathrm{tr}(\mathbf{\Phi}) = \frac{\rho_{\mathrm{p}} C^2\, T_{\mathrm{p}}}{\rho_{\mathrm{p}} C + 1}, \tag{F.8}$$

$$\mathrm{tr}(\mathbf{R}\mathbf{\Phi}) = \frac{\rho_{\mathrm{p}} C^3\, T_{\mathrm{p}}}{\rho_{\mathrm{p}} C + 1}, \tag{F.9}$$

By combining the expressions in (F.2) and (F.5), with the expression in (F.6), the expression in (F.6) can be rewritten as

$$\overline{\mathrm{SINR}}^{\mathrm{BF}} = \frac{\left(\frac{\rho_{\mathrm{p}} C^2\, T_{\mathrm{p}}}{\rho_{\mathrm{p}} C+1}\right)^2}{\frac{1}{\rho_{\mathrm{d}}}\frac{\rho_{\mathrm{p}} C^2\, T_{\mathrm{p}}}{\rho_{\mathrm{p}} C+1} + K\frac{\rho_{\mathrm{p}} C^3\, T_{\mathrm{p}}}{\rho_{\mathrm{p}} C+1}}. \tag{F.10}$$

Using straightforward algebra, expression in (F.10) can also be simplified to

$$\overline{\text{SINR}}^{\text{BF}} = \frac{\frac{\rho_{\text{p}}^2 C^4 T_{\text{p}}}{\rho_{\text{p}} C + 1}}{\frac{1}{\rho_{\text{d}}}\rho_{\text{p}} C^2 + K\rho_{\text{p}} C^3}, \tag{F.11}$$

Multiplying both the nominator and denominator by ρ_{d} and dividing both the nominator and denominator by ρ_{p}, directly leads to

$$\overline{\text{SINR}}^{\text{BF}} = \frac{\frac{\rho_{\text{d}}\rho_{\text{p}} C^4 T_{\text{p}}}{\rho_{\text{p}} C + 1}}{C^2 + KC^3\,\rho_{\text{d}}}, \tag{F.12}$$

Let divide the nominator and denominator by C^4, then expression (F.12) can be simplified to

$$\overline{\text{SINR}}^{\text{BF}} = \frac{\frac{\rho_{\text{d}}\rho_{\text{p}}\, T_{\text{p}}}{\rho_{\text{p}} C + 1}}{\frac{1}{C^2} + \frac{K\rho_{\text{d}}}{C}}, \tag{F.13}$$

Using straightforward algebra, expression in (F.13) can be rewritten as

$$\overline{\text{SINR}}^{\text{BF}} = \frac{\rho_{\text{d}}\rho_{\text{p}}\, T_{\text{p}}}{\left(\frac{1}{C} + K\rho_{\text{d}}\right)\left(\rho_{\text{p}} + \frac{1}{C}\right)}, \tag{F.14}$$

Substituting $C = N/P$ back into expression (F.14), directly leads to expression (F.15), completing the proof.

$$\overline{\text{SINR}}^{\text{BF}} = \frac{\rho_{\text{d}}\rho_{\text{p}} T_{\text{p}}}{\left(P/N + \rho_{\text{p}}\right)\left(P/N + K\rho_{\text{d}}\right)} \tag{F.15}$$

Appendix G

This Appendix provides a proof of the achievable sum rate expression for the BF precoder with P-DoF channel model in (4.42) in the case of high SNR. Recalling the SINR expression of the BF precoder with P-DoF, which is given in (G.1).

$$\overline{\mathrm{SINR}}^{\mathrm{BF}} = \frac{\rho_{\mathrm{d}}\rho_{\mathrm{p}}T_{\mathrm{p}}}{\left(P/N + \rho_{\mathrm{p}}\right)\left(P/N + K\rho_{\mathrm{d}}\right)} \tag{G.1}$$

Let assume that $P/N = D$ for notational convenience. Hence, the expression in (G.1) can be rewritten as

$$\overline{\mathrm{SINR}}^{\mathrm{BF}} = \frac{\rho_{\mathrm{d}}\rho_{\mathrm{p}}T_{\mathrm{p}}}{\left(D + \rho_{\mathrm{p}}\right)\left(D + K\rho_{\mathrm{d}}\right)} \tag{G.2}$$

Let simplify the denominator of (G.2) algebraically as

$$\overline{\mathrm{SINR}}^{\mathrm{BF}} = \frac{\rho_{\mathrm{d}}\rho_{\mathrm{p}}T_{\mathrm{p}}}{\left(D^2 + K\rho_{\mathrm{d}}D + \rho_{\mathrm{p}}D + K\rho_{\mathrm{p}}\rho_{\mathrm{d}}\right)}. \tag{G.3}$$

let divide both the nominator and denominator of expression (G.3) by $\rho_{\mathrm{d}}\rho_{\mathrm{p}}$, thus, (G.3) can be rewritten as

$$\overline{\mathrm{SINR}}^{\mathrm{BF}} = \frac{T_{\mathrm{p}}}{\left(\frac{D^2}{\rho_{\mathrm{d}}\rho_{\mathrm{p}}} + \frac{KD}{\rho_{\mathrm{p}}} + \frac{D}{\rho_{\mathrm{d}}} + K\right)}. \tag{G.4}$$

For the case of high SNR, $\rho_{\mathrm{d}} = \rho_{\mathrm{p}} = \rho \to \infty$, the three left hand side terms in the denominator of (G.4) approach the zero, and hence, the average SINR of the BF precoder can be further simplified to

$$\overline{\mathrm{SINR}}^{\mathrm{BF}} = \left(\frac{T_{\mathrm{p}}}{K}\right). \tag{G.5}$$

Substituting expression (G.5) into (4.26) provides the average achievable sum rate of the BF precoder in (4.42), completing the proof.

Appendix H

To find the optimum training sequence length for BF precoding in a P-DoF channel at high SNR, as given in Proposition 4.3, let first relax the requirement for the training length to be a positive integer and replace T_{p} with a real valued variable $\tau \geq 0$ as shown in (4.44). Treating the high SNR sum rate $\bar{C}^{\mathrm{BF},\infty}$ given in (4.15) as a smooth continuous function of training length τ allows the partial derivative

$$\frac{\partial \bar{C}^{\mathrm{BF},\infty}}{\partial \tau} = -\frac{1}{T_{\mathrm{c}} \ln(2)} \ln\left(1 + \frac{\tau}{K}\right) + \frac{1}{T_{\mathrm{c}} \ln(2)} \left(\frac{T_{\mathrm{c}} - \tau}{\tau + K}\right), \tag{H.1}$$

to be obtained. Setting $\frac{\partial \bar{C}^{\mathrm{BF},\infty}}{\partial \tau} = 0$ yields after some algebraic manipulations the equality

$$\ln\left(1 + \frac{\tau}{K}\right) = \frac{T_{\mathrm{c}} + K}{\tau + K} - 1 = \frac{T_{\mathrm{c}} + K}{K} \left(1 + \frac{\tau}{K}\right)^{-1} - 1. \tag{H.2}$$

Moving the terms containing τ on the left hand side allows (H.2) to be written as

$$\left(\ln\left(1 + \frac{\tau}{K}\right) + 1\right) \exp\left(\ln\left(1 + \frac{\tau}{K}\right) + 1\right) = \frac{T_{\mathrm{c}} + K}{K} \mathrm{e}. \tag{H.3}$$

Clearly the left hand side of (H.3) is of the general form $f(\tau)\,\mathrm{e}^{f(\tau)}$ and can be solved using the Lambert W function [133]. In particular, the Lambert W-function, $W(\cdot)$, satisfies $f(\tau) =$

$W\Big(f(\tau)\, e^{f(\tau)}\Big)$ and as a result, the necessary condition for τ at the stationary point becomes

$$\begin{aligned}\tau &= K \exp\left(W\left(\frac{T_\mathrm{c}+K}{K}\mathrm{e}\right)-1\right)-K \\ &= \frac{K}{\mathrm{e}} \exp \frac{\frac{T_\mathrm{c}+K}{K}\mathrm{e}}{W\left(\frac{T_\mathrm{c}+K}{K}\mathrm{e}\right)}-K, \end{aligned} \tag{H.4}$$

which leads directly to (4.44). The second equality in (H.4) follows from the definition of the Lambert W-function, namely, $W\big(f(\tau)\big)e^{W(f(\tau))} = f(\tau)$.

Finally, since τ is a real valued relaxation of the integer valued training length T_p, the true optimum value is either $\lceil\tau\rceil$ or $\lfloor\tau\rfloor$, i.e., one of the two integers nearest to τ. Combining this with the constraint $T_\mathrm{p} \leq P$ leads to (4.43), completing the proof. The same approach applied to $\bar{C}_\perp^{\mathrm{BF},\infty}$ yields (5.48).

Appendix I

This Appendix describes briefly how the asymptotically tight approximation of the SINR for RZF precoding with an imperfect downlink channel estimation in (4.46) is obtained. The analysis of the SINR for the RZF precoder is given in terms of several auxiliary variables that arise from the asymptotic random matrix theory analysis. In particular, the SINR expression of the RZF precoder under a general correlation model with imperfect channel estimation in (4.46) depending on both the transmit channel covariance matrix $\mathbf{R}$, which appears in the expression (4.37), and the covariance matrix of the MMSE channel estimation through $\mathbf{\Phi}$. The analysis in (4.46) is obtained with the P-DoF channel model given in (2.6) with $\mathbf{R} = \frac{N}{P}\mathbf{A}\mathbf{A}^{\mathrm{H}}$, where $\mathbf{A} \in \mathbb{C}^{N\times P}$ is constructed from $P \leq N$ columns of an arbitrary $N \times N$ unitary matrix so that $\mathbf{A}^{\mathrm{H}}\mathbf{A} = \mathbf{I}_P$.

The trace operator is used to simplify the forms, which equals to the sum of diagonal entries (eigenvalues) of a matrix. It has a property that $\mathrm{tr}(\mathbf{A}^{\mathrm{H}}\mathbf{A}) = \mathrm{tr}(\mathbf{A}\mathbf{A}^{\mathrm{H}})$. With the P-DoF channel model, the channel covariance matrix is given as $\mathrm{tr}(\mathbf{R}) = \frac{N}{P}P$. Additionally, with the P-DoF channel model, the covariance matrix of the MMSE channel estimation can be provided in closed-form as given in the expression (F.5) in Appendix F. Using the aforementioned simplified analyses of the channel covariance matrix and the covariance matrix of the MMSE channel, the expression of the RZF precoder with imperfect channel estimation in (4.30) can be provided in closed-form. The Wolfram Mathematica computer program was also adopted here to simplify the analysis of the RZF precoder.

Bibliography

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 20172022, feb. 2019. [online] available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-738429.pdf," *Cisco*, 2019.

[2] J. Vieira, S. Malkowsky, K. Nieman, Z. Miers, N. Kundargi, L. Liu, I. Wong, V. Öwall, O. Edfors, and F. Tufvesson, "A flexible 100-antenna testbed for massive mimo," in *2014 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2014, pp. 287–293.

[3] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, Feb. 2016.

[4] T. S. Rappaport, W. Roh, and K. Cheun, "Wireless engineers long considered high frequencies worthless for cellular systems. they couldn't be more wrong," *IEEE SPECTRUM*, vol. 51, no. 9, pp. 34–58, Sep. 2014.

[5] E. Dahlman, G. Mildh, S. Parkvall, J. Peisa, J. Sachs, Y. Selén, and J. Sköld, "5G wireless access: requirements and realization," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 42–47, Dec. 2014.

[6] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE journal on selected areas in communications*, vol. 35, no. 6, pp. 1201–1221, 2017.

[7] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE access*, vol. 3, pp. 1206–1232, 2015.

[8] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[9] S. C. Swales, M. A. Beach, D. J. Edwards, and J. P. McGeehan, "The performance enhancement of multibeam adaptive base-station antennas for cellular land mobile radio systems," *IEEE Transactions on Vehicular Technology*, vol. 39, no. 1, pp. 56–67, Feb. 1990.

[10] S. Anderson, B. Hagerman, H. Dam, U. Forssen, J. Karlsson, F. Kronestedt, S. Mazur, and K. J. Molnar, "Adaptive antennas for gsm and tdma systems," *IEEE Personal Communications*, vol. 6, no. 3, pp. 74–86, June 1999.

[11] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[12] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[13] T. L. Marzetta, "Massive mimo: an introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11–22, March 2015.

[14] S. Parkvall, E. Dahlman, A. Furuskar, and M. Frenne, "NR: The new 5G radio access technology," *IEEE Communications Standards Magazine*, vol. 1, no. 4, pp. 24–30, Des., 2017.

[15] T. L. Marzetta, "Blast training: Estimating channel characteristics for high capacity space-time wireless," in *Proceedings. 37th Annual Allerton Conference on Communication, Control, and Computing*, vol. 37, Sept. 1999, pp. 958–966.

[16] J. H. Kotecha and A. M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 546–557, Feb. 2004.

[17] J. Pang, J. Li, L. Zhao, and Z. Lu, "Optimal training sequences for MIMO channel estimation with spatial correlation," in *IEEE Vehicular Technology Conference, Baltimore, MD, USA*, 30 Sept.-3 Oct. 2007, pp. 651–655.

[18] Y. Liu, T. F. Wong, and W. W. Hager, "Training signal design for estimation of correlated MIMO channels with colored interference," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1486–1497, Apr. 2007.

[19] M. Biguesh, S. Gazor, and M. H. Shariat, "Optimal training sequence for mimo wireless systems in colored environments," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3144–3153, Aug. 2009.

[20] E. Bjornson and B. Ottersten, "Training-based bayesian mimo channel and channel norm estimation," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, May 2009, pp. 2701–2704.

[21] E. Björnson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbance," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1807–1820, Mar. 2010.

[22] N. Shariati, J. Wang, and M. Bengtsson, "Robust training sequence design for correlated mimo channel estimation," *IEEE Transactions on Signal Processing*, vol. 62, no. 1, pp. 107–120, Oct. 2013.

[23] N. Shariati, E. Björnson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale mimo with arbitrary statistics," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 815–830, Oct. 2014.

[24] J. Fang, X. Li, H. Li, and F. Gao, "Low-rank covariance-assisted downlink training and channel estimation for FDD massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1935–1947, Mar. 2017.

[25] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. NJ:Prentice-Hall, 1993.

[26] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[27] Björnson, E. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[28] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, Nov. 2016.

[29] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[30] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Journal on selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[31] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[32] S. Wagner, R. Couillet, M. Debbah, and D. T. Slock, "Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4509–4537, July. 2012.

[33] L. Fan, D. Qiao, S. Jin, C.-K. Wen, and M. Matthaiou, "Optimal pilot length for uplink massive MIMO systems with low-resolution ADC," in *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), Rio de Janerio, Brazil*, 10-13 July 2016, pp. 1–5.

[34] R. R. Müller, L. Cottatellucci, and M. Vehkapera, "Blind pilot decontamination," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 773–786, Oct. 2014.

[35] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, "Relative channel reciprocity calibration in MIMO/TDD systems," in *Future Network & Mobile Summit, Florence, Italy*. IEEE, 16-18 June 2010, pp. 1–10.

[36] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Hardware impairments in large-scale MISO systems: Energy efficiency, estimation, and capacity limits," in *18th International Conference on Digital Signal Processing (DSP)*, July July 2013, pp. 1–6.

[37] D. Mi, M. Dianati, L. Zhang, S. Muhaidat, and R. Tafazolli, "Massive mimo performance with imperfect channel reciprocity and channel estimation error," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3734–3749, 2017.

[38] J. Vieira, F. Rusek, O. Edfors, S. Malkowsky, L. Liu, and F. Tufvesson, "Reciprocity calibration for massive MIMO: Proposal, modeling, and validation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3042–3056, 2017.

[39] E. Björnson, J. Hoydis, L. Sanguinetti *et al.*, *Massive MIMO networks: Spectral, energy, and hardware efficiency*. Now Publishers, Inc., 2017, vol. 11, no. 3-4.

[40] Ericsson, "Ericsson mobility report: On the pulse of the network society. technical report 2016. [online] available: https://www.ericsson.com/assets/local/mobility-report/documents/2016/ericsson-mobility-report-november-2016.pdf," *Ericsson*, 2016.

[41] J. Choi, D. J. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 802–814, Oct. 2014.

[42] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, "Pilot beam pattern design for channel estimation in massive MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 787–801, Oct. 2014.

[43] J. So, D. Kim, Y. Lee, and Y. Sung, "Pilot signal design for massive MIMO systems: A received signal-to-noise-ratio-based approach," *IEEE Signal Processing Letters*, vol. 22, no. 5, pp. 549–553, May 2015.

[44] X. Rao and V. K. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3261–3271, June 2014.

[45] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6169–6183, Dec. 2015.

[46] Z. Gao, L. Dai, W. Dai, B. Shim, and Z. Wang, "Structured compressive sensing-based spatio-temporal joint channel estimation for FDD massive MIMO," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 601–617, Feb. 2016.

[47] Y. Han, J. Lee, and D. J. Love, "Compressed sensing-aided downlink channel training for FDD massive MIMO systems," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 2852–2862, Jul. 2017.

[48] Y. Zhou, M. Herdin, A. M. Sayeed, and E. Bonek, "Experimental study of MIMO channel statistics and capacity via the virtual channel representation," in *University Wisconsin, Madison, WI, USA, Technical Report 5*, Feb. 2007, pp. 10–15.

[49] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in *Proceedings of the 18th annual international conference on Mobile computing and networking, Istanbul, Turkey,*. ACM, Aug. 22 - 26, 2012, pp. 53–64.

[50] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3170–3184, April 2016.

[51] A. Adhikary, J. Nam, J. Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: The large-scale array regime," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[52] J. Nam, G. Caire, and J. Ha, "On the role of transmit correlation diversity in multiuser

MIMO systems," *IEEE Transactions on Information Theory*, vol. 63, no. 1, pp. 336–354, Jan 2017.

[53] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niu, "Achievable rates of FDD massive MIMO systems with spatial channel correlation," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2868–2882, May 2015.

[54] Y. Gu and Y. D. Zhang, "Information-theoretic pilot design for downlink channel estimation in FDD massive MIMO systems," *IEEE Transactions on Signal Processing*, 08 March 2019.

[55] B. Tomasi, A. Decurninge, and M. Guillaud, "SNOPS: Short non-orthogonal pilot sequences for downlink channel state estimation in FDD massive MIMO," in *IEEE Globecom Workshops (GC Wkshps), Washington, DC, USA*, 4-8 Dec. 2016, pp. 1–6.

[56] S. Bazzi and W. Xu, "Downlink training sequence design for FDD multiuser massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 65, no. 18, pp. 4732–4744, Sept. 2017.

[57] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[58] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[59] W. Guo and T. O'Farrell, "Capacity-energy-cost tradeoff in small cell networks," in *IEEE 75th Vehicular Technology Conference (VTC Spring)*, 6-9 May 2012, pp. 1–5.

[60] L. Zhao, G. Zhao, and T. O'Farrell, "Efficiency metrics for wireless communications," in *IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 8-11 Sept. 2013, pp. 2825–2829.

[61] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[62] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[63] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[64] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3059–3075, June 2015.

[65] S. Malkowsky, J. Vieira, L. Liu, P. Harris, K. Nieman, N. Kundargi, I. C. Wong, F. Tufvesson, V. Öwall, and O. Edfors, "The world's first real-time testbed for massive mimo: Design, implementation, and validation," *IEEE Access*, vol. 5, pp. 9073–9088, Dec. 2017.

[66] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[67] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE transactions on signal processing*, vol. 54, no. 3, pp. 884–893, Feb. 2006.

[68] ——, "Downlink channel estimation in cellular systems with antenna arrays at base stations using channel probing with feedback," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, no. 9, p. 963649, 2004.

[69] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5g," *IEEE communications magazine*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[70] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-Esposti, H. Hofstetter, P. Kyösti, D. Laurenson, G. Matz *et al.*, "Survey of channel and radio propagation mod-

els for wireless MIMO systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 1, p. 019070, 2007.

[71] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Fortieth Asilomar Conference on Signals, Systems and Computers*. IEEE, 29 Oct.-1 Nov. 2006, pp. 359–363.

[72] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving "massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3226–3239, Sept. 2012.

[73] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[74] H. Huh, A. M. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-complexity scheduling," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2911–2934, May. 2012.

[75] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2845–2866, June 2010.

[76] F. Zhang, S. Sun, Q. Gao, and W. Tang, "Enhanced CSI acquisition for FDD multi-user massive MIMO systems," *IEEE Access*, vol. 6, pp. 23 034–23 042, 23 April 2018.

[77] E. Björnson, E. Jorswieck *et al.*, "Optimal resource allocation in coordinated multi-cell systems," *Foundations and Trends® in Communications and Information Theory*, vol. 9, no. 2–3, pp. 113–381, 2013.

[78] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, Jun. 2003.

[79] H. Weingarten, Y. Steinberg, and S. S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *IEEE transactions on information theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[80] D. Gesbert, M. Kountouris, R. W. Heath, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE signal processing magazine*, vol. 24, no. 5, pp. 36–46, 2007.

[81] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part i: channel inversion and regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, Jan. 2005.

[82] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on selected areas in communications*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[83] A. Wiesel, Y. C. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses." *IEEE Trans. Signal Processing*, vol. 56, no. 9, pp. 4409–4418, Sep. 2008.

[84] R. W. Heath Jr and A. Lozano, *Foundations of MIMO communication*. Cambridge University Press, 2018.

[85] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*. Cambridge University Press, Oct. 2011.

[86] M. H. Hayes, *Statistical digital signal processing and modeling*. John Wiley & Sons, 2009.

[87] S. Stein, "Fading channel issues in system engineering," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 68–89, Feb. 1987.

[88] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, P. Karlsson, and M. Beach, "Modeling of wide-band MIMO radio channels based on NLoS indoor measurements," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 3, pp. 655–665, May 2004.

[89] J. W. Wallace and M. A. Jensen, "Measured characteristics of the MIMO wireless channel," in *IEEE 54th Vehicular Technology Conference. VTC Fall 2001.*, vol. 4, 7-11 Oct. 2001, pp. 2038–2042.

[90] X. Gao, F. Tufvesson, O. Edfors, and F. Rusek, "Measured propagation characteristics for very-large MIMO at 2.6 GHz." in *Asilomar Conference on Signals, Systems, and Computers (ACSCC)*, Nov. 2012, pp. 295–299.

[91] D. Chizhik, G. J. Foschini, M. J. Gans, and R. A. Valenzuela, "Keyholes, correlations, and capacities of multielement transmit and receive antennas," *IEEE Transactions on Wireless Communications*, vol. 1, no. 2, pp. 361–368, April 2002.

[92] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: general model, design criteria, and signal processing," *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 12–25, Nov. 2004.

[93] W. C. Jakes and D. C. Cox, *Microwave mobile communications*. Wiley-IEEE Press, 1994.

[94] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[95] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Transactions on Communications*, vol. 61, no. 6, pp. 2350–2361, June 2013.

[96] H. Q. Ngo, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 22-27 May 2011, pp. 3464–3467.

[97] M. T. Ivrlač, W. Utschick, and J. A. Nossek, "Fading correlations in wireless MIMO communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 819–828, June 2003.

[98] Y.-C. Liang and F. P. S. Chin, "Downlink channel covariance matrix (DCCM) estimation and its applications in wireless DS-CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 2, pp. 222–232, 2001.

[99] B. M. Hochwald and T. Maretta, "Adapting a downlink array from uplink measurements," *IEEE Transactions on Signal Processing*, vol. 49, no. 3, pp. 642–653, March 2001.

[100] X. Mestre, "Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5113–5129, Nov. 2008.

[101] T. L. Marzetta, G. H. Tucci, and S. H. Simon, "A random matrix-theoretic approach to handling singular covariance estimates," *IEEE Transactions on Information Theory*, vol. 57, no. 9, pp. 6256–6271, Sept. 2011.

[102] P. Vallet, P. Loubaton, and X. Mestre, "Improved subspace estimation for multivariate observations of high dimension: the deterministic signals case," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1043–1068, Feb. 2012.

[103] A. Decurninge, M. Guillaud, and D. T. Slock, "Channel covariance estimation in massive MIMO frequency division duplex systems," in *IEEE Globecom Workshops (GC Wkshps), San Diego, CA, USA*, 6–10 Dec. 2015, pp. 1–6.

[104] E. Björnson, L. Sanguinetti, and M. Debbah, "Massive MIMO with imperfect channel covariance information," *2016 50th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA*, pp. 974–978, 6–9 Nov. 2016.

[105] K. Upadhya and S. A. Vorobyov, "Covariance matrix estimation for massive MIMO," *IEEE Signal Processing Letters*, vol. 25, no. 4, pp. 546–550, 2018.

[106] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE Journal on selected areas in communications*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[107] L. You, X. Gao, X.-G. Xia, N. Ma, and Y. Peng, "Pilot reuse for massive MIMO transmission over spatially correlated Rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3352–3366, 2015.

[108] H. Xie, F. Gao, and S. Jin, "An overview of low-rank channel estimation for massive MIMO systems," *IEEE Access*, vol. 4, pp. 7313–7321, 1 Nov. 2016.

[109] P. Stoica and D. A. Linebarger, "Optimization result for constrained beamformer design," *IEEE Signal Processing Letters*, vol. 2, no. 4, pp. 66–67, April 1995.

[110] S. Fiori, A. Uncini, and F. Piazza, "Application of the mec network to principal component analysis and source separation," in *International Conference on Artificial Neural Networks*. Springer, 1997, pp. 571–576.

[111] J. Yang and D. B. Williams, "MIMO transmission subspace tracking with low rate feedback," in *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing.*, vol. 3. IEEE, 23-23 March 2005, pp. iii–405.

[112] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 485–500, April 2016.

[113] A. Cichocki and S.-i. Amari, *Adaptive blind signal and image processing: learning algorithms and applications*. John Wiley & Sons, 2002.

[114] M. Nikpour, J. H. Manton, and G. Hori, "Algorithms on the stiefel manifold for joint diagonalisation," in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2. IEEE, 13-17 May 2002, p. 1481–1484.

[115] F. Liu, C. Masouros, P. V. Amadori, and H. Sun, "An efficient manifold algorithm for constructive interference based constant envelope precoding," *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1542–1546, Oct. 2017.

[116] D. Gabay, "Minimizing a differentiable function over a differential manifold," *Journal of Optimization Theory and Applications*, vol. 37, no. 2, pp. 177–219, 1982.

[117] R. W. Brockett, "Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems," *Linear Algebra and its applications*, vol. 146, pp. 79–91, 1991.

[118] S. T. Smith, "Geometric optimization methods for adaptive filtering," *Ph.D. Thesis, Harvard University, Cambridge, MA*, May 1993.

[119] T. E. Abrudan, J. Eriksson, and V. Koivunen, "Steepest descent algorithms for optimization under unitary matrix constraint," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1134–1147, March 2008.

[120] T. Abrudan, J. Eriksson, and V. Koivunen, "Conjugate gradient algorithm for optimization under unitary matrix constraint," *Signal Processing*, vol. 89, no. 9, pp. 1704–1714, March 2009.

[121] G. W. Stewart, *Matrix Algorithms: Volume 1: Basic Decompositions*. SIAM, 1998.

[122] D. S. Watkins, *Fundamentals of matrix computations*. John Wiley & Sons, 2004, vol. 64.

[123] R. Hunger, *Floating point operations in matrix-vector calculus*. Technische Universität München, Associate Institute for Signal Processing, Tech. Rep. TUM-LNS-TR-05-05, Oct. 2005.

[124] M. Sadeghi, L. Sanguinetti, R. Couillet, and C. Yuen, "Large system analysis of power normalization techniques in massive mimo," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 9005–9017, Oct. 2017.

[125] K. B. Petersen and S. P. Michael, "The matrix cookbook," *University of Denmark, 7*, no. 15, pp. 1704–1714, 2008.

[126] S. Bazzi and W. Xu, "Low-complexity channel estimation in correlated massive MIMO channels," in *WSA 2018; 22nd International ITG Workshop on Smart Antennas, Bochum, Germany*. VDE, 14-16 March 2018, pp. 1–6.

[127] A. Hjrungnes, *Complex-valued matrix derivatives: with applications in signal processing and communications*. Cambridge University Press, 2011.

[128] C. Moler and C. Van Loan, "Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later," *SIAM review*, vol. 45, no. 1, pp. 3–49, 2003.

[129] A. Iserles and A. Zanna, "Efficient computation of the matrix exponential by generalized polar decompositions," *SIAM journal on numerical analysis*, vol. 42, no. 5, pp. 2218–2256, 2005.

[130] L. Armijo, "Minimization of functions having lipschitz continuous first partial derivatives," *Pacific Journal of mathematics*, vol. 16, no. 1, pp. 1–3, 1966.

[131] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 876–890, March 2014.

[132] H. Huh, A. M. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-complexity scheduling," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2911–2934, May. 2012.

[133] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey, and D. E. Knuth, "On the LambertW function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, Dec. 1996.

[134] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Massive MIMO with spatially correlated Rician fading channels," *IEEE Transactions on Communications*, vol. 67, no. 5, pp. 3234–3250, May 2019.

[135] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, March 2017.

[136] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4445–4459, July 2017.

[137] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "How much do downlink pilots improve cell-free massive MIMO?" in *2016 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 4-8 Dec. 2016, pp. 1–7.

[138] E. Björnson and L. Sanguinetti, "Making cell-free massive mimo competitive with mmse processing and centralized implementation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 77–90, 2019.

[139] K. Zheng, L. Zhao, J. Mei, B. Shao, W. Xiang, and L. Hanzo, "Survey of large-scale MIMO systems," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1738–1760, Aug. 2015.

[140] A. Forenza, D. J. Love, and R. W. Heath, "Simplified spatial correlation models for clustered mimo channels with different array configurations," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1924–1934, July 2007.